

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2005, 16:30:06 ; Search time 42 Seconds

(without alignments)
447.894 Million cell updates/sec

Title: US-10-774-076-1

Perfect score: 1308

Sequence: 1 MRAPLPPAVVLSLILGS.....GEAERKKLQENGWHAIA 252

Scoring table:

BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	252	2	US-07-885-089B-7
2	1308	100.0	252	4	US-09-976-594-582
3	1308	100.0	252	4	US-09-919-039-228
4	1308	100.0	322	4	US-09-949-016-11381
5	1177	90.0	226	1	US-07-885-089B-35
6	1015.5	77.6	192	1	US-08-208-008C-9
7	1004.5	76.8	247	2	US-07-885-089B-2
8	1004.5	76.8	247	2	US-07-885-089B-8
9	914	69.9	248	2	US-07-885-089B-6
10	914	69.9	248	2	US-07-885-089B-10
11	874.5	66.9	243	2	US-07-885-089B-4
12	874.5	66.9	243	2	US-07-885-089B-9
13	874.5	66.9	243	6	5202428-2
14	874.5	65.4	243	6	5202428-2
15	855	65.4	158	2	US-07-885-089B-34
16	789	60.3	152	2	US-07-885-089B-36
17	762	58.3	146	2	US-07-885-089B-37
18	700.5	53.6	137	6	5202428-9
19	700.5	53.6	137	6	5202428-9
20	534.5	40.9	147	6	5202428-8
21	534.5	40.9	147	6	5202428-8
22	501	38.3	91	3	US-09-077-977A-1
23	467	35.7	84	6	5202428-10
24	467	35.7	84	6	5202428-10
25	459	35.1	83	2	US-07-885-089B-16
26	454	34.7	87	1	US-07-847-743B-17
27	454	34.7	87	1	US-08-456-201-17

28	454	34.7	87	2	US-08-330-161-15	Sequence 15, Appl
29	454	34.7	87	2	US-08-456-241-17	Sequence 17, Appl
30	454	34.7	87	2	US-08-440-401-15	Sequence 15, Appl
31	454	34.7	87	2	US-08-419-878B-15	Sequence 15, Appl
32	454	34.7	87	3	US-09-173-480-15	Sequence 15, Appl
33	454	34.7	87	5	PCT-US92-04295A-17	Sequence 18, Appl
34	442	33.8	83	2	US-07-885-089B-18	Sequence 18, Appl
35	440	33.6	79	2	US-07-885-089B-33	Sequence 13, Appl
36	433	33.1	85	2	US-07-885-089B-11	Sequence 11, Appl
37	356	27.2	78	2	US-07-885-089B-32	Sequence 32, Appl
38	355	27.1	78	2	US-07-885-089B-30	Sequence 30, Appl
39	347.5	26.6	79	2	US-07-885-089B-31	Sequence 31, Appl
40	320	24.5	87	1	US-07-847-743B-18	Sequence 18, Appl
41	320	24.5	87	1	US-08-456-201-18	Sequence 18, Appl
42	320	24.5	87	2	US-08-330-161-16	Sequence 16, Appl
43	320	24.5	87	2	US-08-456-241-18	Sequence 16, Appl
44	320	24.5	87	2	US-08-440-401-16	Sequence 16, Appl
45	320	24.5	87	2	US-08-419-878B-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-07-885-089B-7
Sequence 7, Application US/0785089B
Patent No. 5830995
GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
APPLICANT: McDonald, Vicki L.
APPLICANT: Bradley, James G.
APPLICANT: Plovman, Gregory D.
TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF
NUMBER OF SEQUENCES: 37
TITLE OF INVENTION: HEPARIN-BINDING EPITHELIAL CELL GROWTH FACTORS
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/885,089B
FILING DATE: 18-MAY-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 5624-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-07-885-089B-7
Query Match 100.0%; Score 1308; DB 2; Length 252;
Best local similarity 100.0%; Pred. No. 8.2e-122; Indels 0; Gaps 0;
Matches 252; Conservative 0; Mismatches 0;
QY 1 MRAPLPPAVVLSLILGSQHYAAGLDNDTYSCKREPSSGDHSDGFEVTSREKSSSG 60

```

Db      1 MRAPLPPAPVVLSTLILSGHYAAGLDLNDTYSGRREPSGDHSDGFEVTSRSMSSG 60
Qy      61 SEISPVSEMPSSSESSGADYDSEYDNEPQIPGYIVDVSVRVQVVKPONTTESENT 120
Db      61 SEISPVSEMPSSSESSGADYDSEYDNEPQIPGYIVDVSVRVQVVKPONTTESENT 120
Qy      121 SDKPRKRRKKGGKNGKRRNRKKKPCNAEFONFCIHGECKYIEHLAVTCCKQOEFGER 180
Db      121 SDKPRKRRKKGGKNGKRRNRKKKPCNAEFONFCIHGECKYIEHLAVTCCKQOEFGER 180
Qy      181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVILTAVAVITVQLRQYVRYKYEGBAERKK 240
Db      181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVILTAVAVITVQLRQYVRYKYEGBAERKK 240
Qy      241 LROENGNVHAIA 252
Db      241 LROENGNVHAIA 252

```

```

RESULT 2
US-09-976-594-582
; Sequence 582, Application US/0976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976, 594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240, 409
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 582
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 3478236CD1
US-09-976-594-582

```

```

Query Match 1
Best Local Similarity 100.0%; Score 1308; DB 4; Length 252;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 MRAPLPPAPVVLSTLILSGHYAAGLDLNDTYSGRREPSGDHSDGFEVTSRSMSSG 60
Db      1 MRAPLPPAPVVLSTLILSGHYAAGLDLNDTYSGRREPSGDHSDGFEVTSRSMSSG 60
Qy      61 SEISPVSEMPSSSESSGADYDSEYDNEPQIPGYIVDVSVRVQVVKPONTTESENT 120
Db      61 SEISPVSEMPSSSESSGADYDSEYDNEPQIPGYIVDVSVRVQVVKPONTTESENT 120
Qy      121 SDKPRKRRKKGGKNGKRRNRKKKPCNAEFONFCIHGECKYIEHLAVTCCKQOEFGER 180
Db      121 SDKPRKRRKKGGKNGKRRNRKKKPCNAEFONFCIHGECKYIEHLAVTCCKQOEFGER 180
Qy      181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVILTAVAVITVQLRQYVRYKYEGBAERKK 240
Db      181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVILTAVAVITVQLRQYVRYKYEGBAERKK 240
Qy      241 LROENGNVHAIA 252
Db      241 LROENGNVHAIA 252

```

```

RESULT 3
US-09-919-039-228
; Sequence 228, Application US/09919039
; Patent No. 6727066

```

```

; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919, 039
; PRIOR FILING DATE: 2002-09-08
; PRIOR APPLICATION NUMBER: 60/222, 113
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 228
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 3478236CD1
US-09-919-039-228

```

```

Query Match 1
Best Local Similarity 100.0%; Score 1308; DB 4; Length 252;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 MRAPLPPAPVVLSTLILSGHYAAGLDLNDTYSGRREPSGDHSDGFEVTSRSMSSG 60
Db      1 MRAPLPPAPVVLSTLILSGHYAAGLDLNDTYSGRREPSGDHSDGFEVTSRSMSSG 60
Qy      61 SEISPVSEMPSSSESSGADYDSEYDNEPQIPGYIVDVSVRVQVVKPONTTESENT 120
Db      61 SEISPVSEMPSSSESSGADYDSEYDNEPQIPGYIVDVSVRVQVVKPONTTESENT 120
Qy      121 SDKPRKRRKKGGKNGKRRNRKKKPCNAEFONFCIHGECKYIEHLAVTCCKQOEFGER 180
Db      121 SDKPRKRRKKGGKNGKRRNRKKKPCNAEFONFCIHGECKYIEHLAVTCCKQOEFGER 180
Qy      181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVILTAVAVITVQLRQYVRYKYEGBAERKK 240
Db      181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVILTAVAVITVQLRQYVRYKYEGBAERKK 240
Qy      241 LROENGNVHAIA 252
Db      241 LROENGNVHAIA 252

```

```

RESULT 4
US-09-949-016-11381
; Sequence 11381, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11381
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11381

```

```

Query Match 1
Best Local Similarity 100.0%; Score 1308; DB 4; Length 322;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MRAPLPAPVYVLSLLILSGHYAAGLDNDYSGKREPFSGDHSADGFEVTSRSEMSG 60
DB 71 MRAPLPAPVYVLSLLILSGHYAAGLDNDYSGKREPFSGDHSADGFEVTSRSEMSG 130
QY 61 SEISPVSEMPSSSPSSGADYDSEYDNEPQIPGYIVDSVREVEQVVKPQNKTESSENT 120
DB 131 SEISPVSEMPSSSPSSGADYDSEYDNEPQIPGYIVDSVREVEQVVKPQNKTESSENT 190
QY 121 SDPRKRRKKGGKGNKRRNKKKPCNAEFONFCIHCECKYIEHLEAVTCKCOQOYFGER 180
DB 191 SDPRKRRKKGGKGNKRRNKKKPCNAEFONFCIHCECKYIEHLEAVTCKCOQOYFGER 250
QY 181 CDEKSMKTHSMIDSSLSKIALAIAAFMSAVILTAIVTQLRQYVRYKGEAEERKK 240
DB 251 CDEKSMKTHSMIDSSLSKIALAIAAFMSAVILTAIVTQLRQYVRYKGEAEERKK 310
QY 241 LROENGVAHAIA 252
DB 311 LROENGVAHAIA 322
RESULT 5
US-07-885-089B-35
; Sequence 35, Application US/07885089B
; Patent No. 5830995
; GENERAL INFORMATION:
; APPLICANT: Shoyab, Mohammed
; APPLICANT: McDonald, Vicki L.
; APPLICANT: Bradley, James G.
; APPLICANT: Plowman, Gregory D.
; TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF
; TITLE OF INVENTION: HEPARIN-BINDING EPITHELIAL CELL GROWTH FACTORS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,089B
; FILING DATE: 18-MAY-1992
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Cortuzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 5624-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 226 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-07-885-089B-35
Query Match 90.0%; Score 1177; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 7.3e-109; Indels 0; Gaps 0;
Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 27 LDNDYTSGRKREPFSGDHSADGFEVTSRSEMSGSEISPVSEMPSSSPSSGADYDSE 86
DB 1 LDNDYTSGRKREPFSGDHSADGFEVTSRSEMSGSEISPVSEMPSSSPSSGADYDSE 60

QY 87 YDNEPQIPGYIVDSVREVEQVVKPQNKTESSENTSDPKRRKKKKGGKGNKRRNKKKNPC 146
DB 61 YDNEPQIPGYIVDSVREVEQVVKPQNKTESSENTSDPKRRKKKKGGKGNKRRNKKKNPC 120
QY 147 NAEFONFCIHCECKYIEHLEAVTCKCOQOYFGERCGEKSMKTHSMIDSSLSKIALAIAA 206
DB 121 NAEFONFCIHCECKYIEHLEAVTCKCOQOYFGERCGEKSMKTHSMIDSSLSKIALAIAA 180
QY 207 FMSAVILTAIVTQLRQYVRYKGEAEERKKLROENGVAHAIA 252
DB 181 FMSAVILTAIVTQLRQYVRYKGEAEERKKLROENGVAHAIA 226
RESULT 6
US-08-208-008C-9
; Sequence 9, Application US/08208008C
; Patent No. 5631147
; GENERAL INFORMATION:
; APPLICANT: MEISSNER, ET AL.
; TITLE OF INVENTION: Transforming Growth Factor
; TITLE OF INVENTION: Alpha - H1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/208,008C
; FILING DATE: March 8, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-98
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-208-008C-9
Query Match 77.6%; Score 1015.5; DB 1; Length 192;
Best Local Similarity 99.5%; Pred. No. 6.5e-93; Indels 1; Gaps 1;
Matches 192; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MRAPLPAPVYVLSLLILSGHYAAGLDNDYSGKREPFSGDHSADGFEVTSRSEMSG 60
DB 1 MRAPLPAPVYVLSLLILSGHYAAGLDN-TSGKREPFSGDHSADGFEVTSRSEMSG 59
QY 61 SEISPVSEMPSSSPSSGADYDSEYDNEPQIPGYIVDSVREVEQVVKPQNKTESSENT 120
DB 60 SEISPVSEMPSSSPSSGADYDSEYDNEPQIPGYIVDSVREVEQVVKPQNKTESSENT 119
QY 121 SDPRKRRKKGGKGNKRRNKKKPCNAEFONFCIHCECKYIEHLEAVTCKCOQOYFGER 180

Db 120 SDPKRRKKGGKNGKRRNRKKKPCNAEFQNFCTHGECKYTEHLEAVTCKCOQYFGER 179
Qy 181 CGEKSMTKTHSMID 193
Db 180 CGEKSMTKTHSMID 192

RESULT 7
US-07-885-089B-2
Sequence 2, Application US/07885089B
Patent No. 5830995
GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
APPLICANT: McDonald, Vicki L.
APPLICANT: Bradley, James G.
APPLICANT: Plowman, Gregory D.
TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF
NUMBER OF SEQUENCES: 37
HEPARIN-BINDING EPITHELIAL CELL GROWTH FACTORS
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/885,089B
FILING DATE: 18-MAY-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 5624-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-885-089B-2

Query Match 76.8%; Score 1004.5; DB 2; Length 247;
Best Local Similarity 77.5%; Pred. No. 1.2e-91;
Matches 196; Conservative 19; Mismatches 31; Indels 7; Gaps 2;
Qy 1 MRAPLLPAPVYVLLILGSGHYAAGLDNDTVSGKREPPSGHSDGFEVTSRSMSSG 60
Db 1 MRAPLLPAPVYVLLILGSGHYAAGLDNDTVSGKREPPSGHSDGFEVTSRSMSSG 60
Qy 61 SEISPVSEMPSSSPSSGADYDSEYDNEPQIPGYIVDSVRAVEQVVPONKTESERT 120
Db 61 SEISPVSEMPSSSPSSGADYDSEYDNEPQIPGYIVDSVRAVEQVVPONKTESERT 120
Qy 121 SDPKRRKKGGKNGKRRNRKKKPCNAEFQNFCTHGECKYTEHLEAVTCKCOQYFGER 180
Db 121 SDPKRRKKGGKNGKRRNRKKKPCNAEFQNFCTHGECKYTEHLEAVTCKCOQYFGER 180
Qy 181 CGEKSMTKTHSMIDSSISKTALAAIAAFSAVILTVAV-ITVQLRQYVRKYEGEAERK 239
Db 181 CGEKSMTKTHSMIDSSISKTALAAIAAFSAVILTVAV-ITVQLRQYVRKYEGEAERK 239
Qy 175 CGEKSMTKTHSMIDSSISKTALAAIAAFSAVILTVAV-ITVQLRQYVRKYEGEAERK 234
Db 175 CGEKSMTKTHSMIDSSISKTALAAIAAFSAVILTVAV-ITVQLRQYVRKYEGEAERK 234
Qy 240 KLRQENGVAHA 252
Db 240 KLRQENGVAHA 252

Db 235 KLRQENGVAHA 247

RESULT 8
US-07-885-089B-8
Sequence 8, Application US/07885089B
Patent No. 5830995
GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
APPLICANT: McDonald, Vicki L.
APPLICANT: Bradley, James G.
APPLICANT: Plowman, Gregory D.
TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF
NUMBER OF SEQUENCES: 37
HEPARIN-BINDING EPITHELIAL CELL GROWTH FACTORS
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/885,089B
FILING DATE: 18-MAY-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 5624-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-07-885-089B-8

Query Match 76.8%; Score 1004.5; DB 2; Length 247;
Best Local Similarity 77.5%; Pred. No. 1.2e-91;
Matches 196; Conservative 19; Mismatches 31; Indels 7; Gaps 2;
Qy 1 MRAPLLPAPVYVLLILGSGHYAAGLDNDTVSGKREPPSGHSDGFEVTSRSMSSG 60
Db 1 MRAPLLPAPVYVLLILGSGHYAAGLDNDTVSGKREPPSGHSDGFEVTSRSMSSG 60
Qy 61 SEISPVSEMPSSSPSSGADYDSEYDNEPQIPGYIVDSVRAVEQVVPONKTESERT 120
Db 61 SEISPVSEMPSSSPSSGADYDSEYDNEPQIPGYIVDSVRAVEQVVPONKTESERT 120
Qy 121 SDPKRRKKGGKNGKRRNRKKKPCNAEFQNFCTHGECKYTEHLEAVTCKCOQYFGER 180
Db 121 SDPKRRKKGGKNGKRRNRKKKPCNAEFQNFCTHGECKYTEHLEAVTCKCOQYFGER 180
Qy 181 CGEKSMTKTHSMIDSSISKTALAAIAAFSAVILTVAV-ITVQLRQYVRKYEGEAERK 239
Db 181 CGEKSMTKTHSMIDSSISKTALAAIAAFSAVILTVAV-ITVQLRQYVRKYEGEAERK 239
Qy 175 CGEKSMTKTHSMIDSSISKTALAAIAAFSAVILTVAV-ITVQLRQYVRKYEGEAERK 234
Db 175 CGEKSMTKTHSMIDSSISKTALAAIAAFSAVILTVAV-ITVQLRQYVRKYEGEAERK 234
Qy 240 KLRQENGVAHA 252
Db 240 KLRQENGVAHA 252

```

1 GENERAL INFORMATION:
2 APPLICANT: Shoyab, Mohammed
3 APPLICANT: McDonald, Vicki L.
4 APPLICANT: Bradley, James G.
5 APPLICANT: Plowman, Gregory D.
6 TITLE OF INVENTION: AMPHIREGLINS: A FAMILY OF
7 TITLE OF INVENTION: HEPARIN-BINDING EPITHELIAL CELL GROWTH FACTORS
8 NUMBER OF SEQUENCES: 37
9 CORRESPONDENCE ADDRESS:
10 ADDRESSEE: Pennie & Edmonds LLP
11 STREET: 1155 Avenue of the Americas
12 CITY: New York
13 STATE: New York
14 COUNTRY: USA
15 ZIP: 10036
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 COMPUTER: IBM PC compatible
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: PatentIn Release #1.0, Version #1.30
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/07/885,089B
23 FILING DATE: 18-MAY-1992
24 CLASSIFICATION: 530
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Coruzzi, Laura A.
27 REGISTRATION NUMBER: 30,742
28 REFERENCE/DOCKET NUMBER: 5624-174
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 212-790-9090
31 TELEFAX: 212-869-9741
32 TELEX: 66141 PENNIE
33 INFORMATION FOR SEQ ID NO: 10:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 248 amino acids
36 TYPE: amino acid
37 STRANDEDNESS: single
38 TOPOLOGY: unknown
39 MOLECULE TYPE: protein
40 US-07-885-089B-10
41
42 Query Match 69.9%; Score 914; DB 2; Length 248;
43 Basic Local Similarity 69.0%; Pred. No. 1,je-82;
44 Matches 176; Conservative 28; Mismatches 41; Indels 10; Gaps 3;
45
46 Db 1 MRAPLLPAPVAVLSLLILGSGHYAAGDLNDPTYSGKKEPSPGDSADGFEVTSRSEMSG 60
47 1 MRTLPLPLASVLLILVLVSGHYAALALINDPSSGKESLSGDHSAGL-----ELSVG 54
48 61 SEISPVSEMPSPSEPPSGADYDYSEEDYNEPQIPGYIVDSVVRVEQYVKKPDPNTESENT 120
49 55 REVSTISEMPSPGSELSTG-DYDYSEEDYNEPQISGYIIDSVRVEQYIKPKNKTTEGKS 113
50
51 Oy 121 SDPRKRRKKGGKGNKRRNRKKKQPCNAEONFCIHCEKTYIELLNAVTCQOQRYGER 180
52 Db 114 TTKRRKKRRKKGGKGBGRNKKQPCNAKQNFCHGECRYIELLEVTCNCHQDYGER 173
53
54 Oy 181 CGEKSMKTHSMINSLSKIALAIAFMAFVAIVLFAV---VITVQLROYVRKYEGRAE 237
55 Db 174 CGEKSMKTHSEDDKSLIAVAVVTIVSAIIILAIGIVITVHLMKRIFREYEGTEE 233
56
57 Oy 238 RKGLRQENGVAIA 252
58 Db 234 RRRLRQENGVAIA 248
59
60 RESULT 11
61 US-07-885-089B-4
62 Sequence 4, Application US/07885089B
63 Patent No. 5830995
64 GENERAL INFORMATION:
65 APPLICANT: Shoyab, Mohammed
66 APPLICANT: McDonald, Vicki L.

```

APPLICANT: Bradley, James G.
APPLICANT: Plozman, Gregory D.
TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/885,089B
FILING DATE: 18-MAY-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 5624-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 243 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-885-089B-4

Query Match 66.9%; Score 874.5; DB 2; Length 243;
Best Local Similarity 69.7%; Pred. No. 9.3e-79;
Matches 177; Conservative 24; Mismatches 40; Indels 13; Gaps 5;

QY 1 MRAPLPAPAVVLTLLIGSGHYAAGLDLNDYSGKREPFSGDSADGFEVTSRSEMSG 60
DB 1 MRPPLSLALVSLVLSLVGSHYAGLELNGTSSGKGPSSGDSAGLVV----- 51
QY 61 SEISPVSEMPSSSESGADYDYSEEDNEPQIPGYIVDSVRVEOVYKPPONTSESENT 120
DB 52 SEVSTISEMPSSSELSGTG-DYDYSEEDNEPQISGIYVDSVRVEOVYKPPONTSESENT 110
QY 121 SDKPRKKKKGGKGNKRNRR-KKKNPCNAEPONFCIHGECKYIEHLBAVTCCKQOEYFGE 179
DB 111 SEKPRKKKKGGKGNKRNRR-KKKNPCNAEPONFCIHGECKYIEHLBAVTCCKQOEYFGE 170
QY 180 RCGESKMTHTMISLSKIALAIAAFMSAVILITAVAVT-VOLROYVRYKGEBAER 238
DB 171 RCGEKMTKQKDDSDSKIALAIAIIVFSAVSVAIGIITVNLKRFREYE-EAER 229
QY 239 KCLRQENGVAIA 252
DB 230 RLRLQENGTAIAIA 243

RESULT 12
US-07-885-089B-9
Sequence 9, Application US/07885089B
Patent No. 5830995
GENERAL INFORMATION:
APPLICANT: Shoyab, Mohammed
APPLICANT: McDonald, Vicki L.
APPLICANT: Bradley, James G.
APPLICANT: Plozman, Gregory D.
TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF
TITLE OF INVENTION: HEPARIN-BINDING EPITHELIAL CELL GROWTH FACTORS

NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/885,089B
FILING DATE: 18-MAY-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 5624-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 243 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-07-885-089B-9

Query Match 66.9%; Score 874.5; DB 2; Length 243;
Best Local Similarity 69.7%; Pred. No. 9.3e-79;
Matches 177; Conservative 24; Mismatches 40; Indels 13; Gaps 5;

QY 1 MRAPLPAPAVVLTLLIGSGHYAAGLDLNDYSGKREPFSGDSADGFEVTSRSEMSG 60
DB 1 MRPPLSLALVSLVLSLVGSHYAGLELNGTSSGKGPSSGDSAGLVV----- 51
QY 61 SEISPVSEMPSSSESGADYDYSEEDNEPQIPGYIVDSVRVEOVYKPPONTSESENT 120
DB 52 SEVSTISEMPSSSELSGTG-DYDYSEEDNEPQISGIYVDSVRVEOVYKPPONTSESENT 110
QY 121 SDKPRKKKKGGKGNKRNRR-KKKNPCNAEPONFCIHGECKYIEHLBAVTCCKQOEYFGE 179
DB 111 SEKPRKKKKGGKGNKRNRR-KKKNPCNAEPONFCIHGECKYIEHLBAVTCCKQOEYFGE 170
QY 180 RCGESKMTHTMISLSKIALAIAAFMSAVILITAVAVT-VOLROYVRYKGEBAER 238
DB 171 RCGEKMTKQKDDSDSKIALAIAIIVFSAVSVAIGIITVNLKRFREYE-EAER 229
QY 239 KCLRQENGVAIA 252
DB 230 RLRLQENGTAIAIA 243

RESULT 13
5202428-2
Patent No. 5202428
TITLE OF INVENTION: DNA ENCODING NEUTROTROPHIC GROWTH FACTOR
NUMBER OF SEQUENCES: 10
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/590,359
FILING DATE: 27-SEP-1990
Prior APPLICATION DATA:
APPLICATION NUMBER: 541,276
FILING DATE: 20-JUN-1990
SEQ ID NO: 2;
LENGTH: 243

5202428-2

Query Match

66.9%; Score 874.5; DB 6; Length 243;

Best Local Similarity 69.7%; Pred. No. 9.3e-79;

Matches 177; Conservative 24; Mismatches 40; Indels 13; Gaps 5;

QY 1 MRAPLLPAPVVLILLIISGHHYAAGLDLNDYSGKREPFSGDHSADGFEVTSRSEMSG 60

DB 1 MRPSTSLALSVLVLVSGHHYAAGLELNGTSSGKGPSSGDSAGGLV----- 51

QY 61 SEISPVSEMPSSSGADYDYSSEYDNEPQIPGYIVDSVRVEQVVKPQNTSESENT 120

DB 52 SEVETISEMPSSGSLSTG-DYDYSSEYDNEPQISGYIVDSVRVEQVVKPQNTSESENT 110

QY 121 SDPRKRRKKGGKGNKRRNR-KKQPCNAEFQNTCHGECKYIEHLAVTCCKQOEYFGE 179

DB 111 SEKPRKRRKKGGKGNKRRNR-KKQPCNAEFQNTCHGECKYIEHLAVTCCKQOEYFGE 170

QY 180 RCGEKMKTQKKDDSKLALAIIVFSAVSAIIGITAVLLRKRFRFRE-EABER 238

DB 171 RCGEKMKTQKKDDSKLALAIIVFSAVSAIIGITAVLLRKRFRFRE-EABER 229

QY 239 KGLRQENGVAIA 252

DB 230 RLRLQENGTAIA 243

RESULT 14

5202428-2

Patent No. 5202428

APPLICANT: SCHUBERT, DAVID

TITLE OF INVENTION: DNA ENCODING NEUROTROPIC GROWTH FACTOR

NUMBER OF SEQUENCES: 10

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/590,359

FILING DATE: 27-SEP-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 541,276

FILING DATE: 20-JUN-1990

SEQ ID NO:2

LENGTH: 243

5202428-2

Query Match

66.9%; Score 874.5; DB 6; Length 243;

Best Local Similarity 69.7%; Pred. No. 9.3e-79;

Matches 177; Conservative 24; Mismatches 40; Indels 13; Gaps 5;

QY 1 MRAPLLPAPVVLILLIISGHHYAAGLDLNDYSGKREPFSGDHSADGFEVTSRSEMSG 60

DB 1 MRPSTSLALSVLVLVSGHHYAAGLELNGTSSGKGPSSGDSAGGLV----- 51

QY 61 SEISPVSEMPSSSGADYDYSSEYDNEPQIPGYIVDSVRVEQVVKPQNTSESENT 120

DB 52 SEVETISEMPSSGSLSTG-DYDYSSEYDNEPQISGYIVDSVRVEQVVKPQNTSESENT 110

QY 121 SDPRKRRKKGGKGNKRRNR-KKQPCNAEFQNTCHGECKYIEHLAVTCCKQOEYFGE 179

DB 111 SEKPRKRRKKGGKGNKRRNR-KKQPCNAEFQNTCHGECKYIEHLAVTCCKQOEYFGE 170

QY 180 RCGEKMKTQKKDDSKLALAIIVFSAVSAIIGITAVLLRKRFRFRE-EABER 238

DB 171 RCGEKMKTQKKDDSKLALAIIVFSAVSAIIGITAVLLRKRFRFRE-EABER 229

QY 239 KGLRQENGVAIA 252

DB 230 RLRLQENGTAIA 243

RESULT 15

US-07-885-089B-34

Sequence 34, Application US/07885089B

Patent No. 5830995

GENERAL INFORMATION:

APPLICANT: Shoyab, Mohammed

APPLICANT: McDonald, Vicki L.

APPLICANT: Bradley, James G.

APPLICANT: Plowman, Gregory D.

TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESS: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/885,089B

FILING DATE: 18-MAY-1992

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Cortuzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 5624-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-9741

TELEX: 66141 PENNIB

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:

LENGTH: 158 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULAR TYPE: protein

US-07-885-089B-34

Query Match

65.4%; Score 855; DB 2; Length 158;

Best Local Similarity 100.0%; Pred. No. 4.4e-77;

Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 LDNDYSGKREPFSGDHSADGFEVTSRSEMSGSEISPVSEMPSSSGADYDYSER 86

DB 1 LDNDYSGKREPFSGDHSADGFEVTSRSEMSGSEISPVSEMPSSSGADYDYSER 60

QY 87 YDNEPQIPGYIVDSVRVEQVVKPQNTSESENTSDKPKKKKKGNKRRNRKKKNPC 146

DB 61 YDNEPQIPGYIVDSVRVEQVVKPQNTSESENTSDKPKKKKKGNKRRNRKKKNPC 120

QY 147 NAEFQNTCHGECKYIEHLAVTCCKQOEYFGERCGEK 184

DB 121 NAEFQNTCHGECKYIEHLAVTCCKQOEYFGERCGEK 158

Search completed: February 26, 2005, 16:45:48

Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2005, 15:47:23 ; Search time 164 Seconds

(without alignments)
594.290 Million cell updates/sec

Title: US-10-774-076-1

Perfect score: 1308

Sequence: 1 MRAPLPPAPVLSLILGS.....GEAEKRLKQENGNAHA 252

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692.

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	252	6 ABR47391	ABR47391 Breast ca
2	1308	100.0	252	6 ABUS7614	ABUS7614 Different
3	1308	100.0	252	8 ADE77063	ADH17052 Human pro
4	1308	100.0	252	8 ADH17052	ADH17052 Human amp
5	1308	100.0	252	8 ADL12853	ADL12853 Human ste
6	1308	100.0	252	8 ADJ75567	ADJ75567 Marker ge
7	1308	100.0	252	8 ADP07715	ADP07715 Human sec
8	1308	100.0	252	8 ADR73587	ADR73587 wild-type
9	1308	100.0	252	8 ADR67228	ADR67228 Human bla
10	1308	100.0	252	4 AAG75109	AAG75109 Human col
11	1295	99.0	235	7 AAP95447	AAP95447 Amphiregu
12	1226	93.7	235	7 ADC71538	ADC71538 Human NOV
13	1226	93.7	235	8 ADN33949	ADN33949 Human nov
14	1177	90.0	226	2 AAW73157	AAW73157 Human amp
15	1004.5	76.8	247	2 AAW73160	AAW73160 Bovine am
16	918	70.2	248	8 ADJ76293	ADJ76293 Marker ge
17	914	69.9	248	2 AAW73162	AAW73162 Mouse amp
18	874.5	66.9	243	2 AAR34705	AAR34705 J5C1 prec
19	874.5	66.9	243	2 AAW73161	AAW73161 Rat amphi
20	855	65.4	158	2 AAW73156	AAW73156 Soluble H
21	789	60.3	152	2 AAW73158	AAW73158 Human amp
22	762	58.3	146	2 AAW73159	AAW73159 Human amp
23	501	38.3	91	2 AAW37304	AAW37304 Amphiregu
24	462	35.3	84	1 AAP90449	AAP90449 Sequence
25	459	35.1	82	2 AAW73153	AAW73153 Human amp

26	442	33.8	82	2 AAW73154	AAW73154 Human amp
27	440	33.6	107	1 AAP95449	AAP95449 Sequence
28	356	27.2	78	2 AAW73151	AAW73151 Mouse amp
29	355	27.1	78	2 AAW73149	AAW73149 Bovine am
30	347.5	26.6	79	2 AAW73150	AAW73150 Rat amphi
31	325.5	24.9	82	1 AAP95451	AAP95451 Protein s
32	283	21.6	49	8 ADN48882	ADN48882 Amphiregu
33	259	19.8	44	6 AAE36798	AAE36798 Human amp
34	246.5	18.8	42	2 AAR31735	AAR31735 Heregulin
35	231	17.7	46	5 AAG66053	AAG66053 Mouse amp
36	216	16.5	46	3 AAB34955	AAB34955 Gene 29 h
37	211.5	16.2	208	2 AAR33998	AAR33998 EGF/HB-EH
38	211.5	16.2	208	2 AAR61390	AAR61390 Diphtheri
39	211.5	16.2	208	2 AAR80785	AAR80785 Human pre
40	211.5	16.2	208	2 AAR80787	AAR80787 Monkey pr
41	211.5	16.2	208	2 AAR32898	AAR32898 Monkey HB
42	211.5	16.2	208	2 AAR32897	AAR32897 Human HB
43	211.5	16.2	208	3 AAY58999	AAY58999 Human hep
44	211.5	16.2	208	3 AAY90406	AAY90406 Human HBE
45	211.5	16.2	208	3 AAY90408	AAY90408 Monkey HB

ALIGNMENTS

RESULT 1	
ID ABR47391	standard; protein; 252 AA.
XX	
AC ABR47391;	
XX	
DT 12-JUN-2003	(first entry)
XX	
DE Breast cancer associated protein sequence SEQ ID NO:13.	
XX	
KM Human, breast cancer; cytostatic; gene therapy.	
XX	
OS Homo sapiens.	
XX	
PN WC0200304989-A2.	
XX	
PD 16-JAN-2003.	
XX	
PF 21-JUN-2002; 2002W0-US019669.	
XX	
PR 21-JUN-2001; 2001US-0299887P.	
PR 27-JUN-2001; 2001US-0301572P.	
PR 18-JUL-2001; 2001US-0306501P.	
PR 25-SEP-2001; 2001US-0325002P.	
PR 05-MAR-2002; 2002US-0362585P.	
PR 14-MAY-2002; 2002US-0380391P.	
XX	
PA (MIL-) MILENITUM PHARM INC.	
PI Lillie J, Ganavayaru M, Glatz K, Hoerh S, Kamatkar S;	
PI Werten M, Monahan DE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;	
PI Best RC, Horibagyi GN, Puztai L, Meric F, Sahin A, Mills GB;	
XX	
DR N-PSDB; ACC50082.	
XX	
PT WPI; 2003-210381/20.	
XX	
PT Breast cancer diagnosis or treatment by comparing the level of expression	
PT of a marker in a patient sample with that in the control non-breast	
PT cancer sample.	
XX	
PS Claim 1; SEQ ID NO 13; 128bp; English.	
XX	
CC The present invention describes a method for assessing whether a patient	
CC is afflicted with breast cancer. The method comprises comparing the level	
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and	
CC ABR47386 to ABR47632) in a patient sample and the normal level of	
CC expression of the marker in a control non-breast cancer sample, where a	
CC significant increase in the level of expression of the marker in the	

CC patient sample and the normal level is an indication that the patient is
 CC afflicted with breast cancer. The breast cancer associated sequences from
 CC the present invention have cytostatic activities and can be used in gene
 CC therapy. The method is useful for diagnosing and treating breast cancer.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences
 SO Sequence 252 AA;

Query Match
 Best Local Similarity 100.0%; Score 1308; DB 6; Length 252;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAPLLPAPVVLSTLLIGSGHYAAGLDNDTYSGRKEPPSGHSDAGFEVTSRSEMSG 60
 DB 1 MRAPLLPAPVVLSTLLIGSGHYAAGLDNDTYSGRKEPPSGHSDAGFEVTSRSEMSG 60
 QY 61 SEISPVSEMPSSSEPSGADYDSEYDNEPQIPGYIVDVSVRVQVVPKPKTESSENT 120
 DB 61 SEISPVSEMPSSSEPSGADYDSEYDNEPQIPGYIVDVSVRVQVVPKPKTESSENT 120
 QY 121 SDKPRKKKKGGKNGKRRNRKKKPCNAEFONFCIHGECKTIEHLEAVTCCKQOEFYFGER 180
 DB 121 SDKPRKKKKGGKNGKRRNRKKKPCNAEFONFCIHGECKTIEHLEAVTCCKQOEFYFGER 180
 QY 181 CGEKSMTKSHMIDSLSKIALAIAAFMSAVITLTAVALITVQLRQVYRKYEGBAERKK 240
 DB 181 CGEKSMTKSHMIDSLSKIALAIAAFMSAVITLTAVALITVQLRQVYRKYEGBAERKK 240
 QY 241 LRQENGNVHAIYA 252
 DB 241 LRQENGNVHAIYA 252

RESULT 2
 ID ABUS7614
 XX ABUS7614 standard; protein; 252 AA.

AC ABUS7614;
 DT 09-APR-2003. (first entry)

DE Differentially expressed breast cancer associated protein #1.
 XX Breast cancer; differential gene expression; BC-cDNA;
 KW breast cancer treatment; breast cancer monitoring;
 KW breast cancer treatment; breast cancer staging.
 OS Homo sapiens.

XX US2002156263-A1.

PD 24-OCT-2002.

PF 04-OCT-2001; 2001US-00974298.

PR 05-OCT-2000; 2000US-0238331P.

PA (CHEN/) CHEN H.

PI Chen H;

DR WPI; 2003-182653/18.

XX New cDNAs, which are differentially expressed in (metastatic) breast
 PT cancer useful for diagnosing or staging, breast cancer, or for monitoring
 XX the treatment of breast cancer in an individual.
 XX Example; SEQ ID NO 2; 30pp; English.

CC The invention describes a combination of cDNAs (designated BC-cDNAs),
 CC which are differentially expressed in breast cancer. The combination

CC includes 152 cDNA sequences, or their complements. The protein encoded by
 CC any of these BC-cDNAs is useful for screening several molecules or
 CC compounds to identify at least one ligand that specifically binds the
 CC protein, producing or preparing polyclonal or monoclonal antibodies, or
 CC purifying antibodies from a sample. The antibodies, which specifically
 CC bind the protein differentially expressed in breast cancer, are useful
 CC for detecting the expression of a protein in a sample. The BC-cDNAs are also
 CC useful for diagnosing, monitoring the treatment of, or staging, breast
 CC cancer. This is the amino acid sequence of a differentially expressed
 CC breast cancer associated protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the US patent office at
 CC segdata.uspto.gov/sequence.html?docid=20020156263
 SO Sequence 252 AA;

Query Match
 Best Local Similarity 100.0%; Score 1308; DB 6; Length 252;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAPLLPAPVVLSTLLIGSGHYAAGLDNDTYSGRKEPPSGHSDAGFEVTSRSEMSG 60
 DB 1 MRAPLLPAPVVLSTLLIGSGHYAAGLDNDTYSGRKEPPSGHSDAGFEVTSRSEMSG 60
 QY 61 SEISPVSEMPSSSEPSGADYDSEYDNEPQIPGYIVDVSVRVQVVPKPKTESSENT 120
 DB 61 SEISPVSEMPSSSEPSGADYDSEYDNEPQIPGYIVDVSVRVQVVPKPKTESSENT 120
 QY 121 SDKPRKKKKGGKNGKRRNRKKKPCNAEFONFCIHGECKTIEHLEAVTCCKQOEFYFGER 180
 DB 121 SDKPRKKKKGGKNGKRRNRKKKPCNAEFONFCIHGECKTIEHLEAVTCCKQOEFYFGER 180
 QY 181 CGEKSMTKSHMIDSLSKIALAIAAFMSAVITLTAVALITVQLRQVYRKYEGBAERKK 240
 DB 181 CGEKSMTKSHMIDSLSKIALAIAAFMSAVITLTAVALITVQLRQVYRKYEGBAERKK 240
 QY 241 LRQENGNVHAIYA 252
 DB 241 LRQENGNVHAIYA 252

RESULT 3
 ID ADE77063
 XX ADE77063 standard; protein; 252 AA.

AC ADE77063;

DT 29-JAN-2004 (first entry)

DE Human protein expressed in a liver disorder #62.

XX human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
 KW tumour; liver; inflammatory disorder; immune response disorder;
 KW high-throughput screening; differential gene expression; gene therapy.
 OS Homo sapiens.

XX US2003108871-A1.

PD 12-JUN-2003.

PF 30-JUL-2001; 2001US-00919039.

PR 28-JUL-2000; 2000US-0222113P.

PA (KASER/) KASER M R.

PI Kaser MR;

DR WPI; 2004-031227/03.
 DR N-PSDB; ADE77062.

XX Composition comprising several cDNAs that are differentially expressed in

PT treated human C3A liver cell cultures, useful for treating liver
disorders.
XX
PS Claim 1, SEQ ID NO 228; 41bp, English.
XX
CC The invention relates to a composition comprising several cDNAs that are
differentially expressed in a liver disorder. The composition is useful
for treating liver disorder such as hyperlipidemia, hypertension, type
II diabetes, tumours of the liver and disorders of the inflammatory and
immune response. The composition is useful for a high-throughput method
of screening several molecules or compounds to identify a ligand which
specifically binds a cDNA. A protein encoded by the cDNA is useful for a
high-throughput method for using a protein to screen several molecules or
compounds to identify at least one ligand which specifically binds the
protein which involves combining the protein encoded by the cDNA with
several of molecules or compounds under conditions to allow specific
binding, and detecting specific binding between the protein and a
molecule or compound, therefore identifying a ligand which specifically
binds the protein. The composition is useful for detecting and
quantifying differential gene expression, can be used in gene therapy, to
formulate prognosis and to design a treatment regimen and to monitor the
efficacy of treatment. The present sequence represents the amino acid
sequence of a protein encoded by a cDNA differentially expressed in a
liver disorder.
XX
SQ Sequence 252 AA:
Query Match 100.0%; Score 1308; DB 8; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.2e-113;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRAPLPPAPVYVSLILGSGHYAAGLDNDYSGKREPFGSDHSDGFVTSRSESSG 60
DB 1 MRAPLPPAPVYVSLILGSGHYAAGLDNDYSGKREPFGSDHSDGFVTSRSESSG 60
QY 61 SEISPVSEMPSSSPSGADYDSEEDNEPQIPGYIVDDSVRVEQVVKPQNTSESENT 120
DB 61 SEISPVSEMPSSSPSGADYDSEEDNEPQIPGYIVDDSVRVEQVVKPQNTSESENT 120
QY 121 SDRPKRRKKGGKGNKRRNRKKNPCNAEFONFCIHGCKYIHLAVTCCKQOEYFGER 180
DB 121 SDRPKRRKKGGKGNKRRNRKKNPCNAEFONFCIHGCKYIHLAVTCCKQOEYFGER 180
QY 181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVITVAVITVQLRQYVRKTEGAEERKK 240
DB 181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVITVAVITVQLRQYVRKTEGAEERKK 240
QY 241 LROENGNVHAIA 252
DB 241 LROENGNVHAIA 252
RESULT 4
ADL17052
ID ADL17052 standard; protein; 252 AA.
XX
AC ADL17052;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human amphiregulin protein.
XX
KM tyrosine kinase activity; type 1 plasminogen activator inhibitor; PAI-1;
KM TIMP-1; tissue inhibitor of metalloproteinase 1; vinculin;
KM vascular endothelial growth factor; VEGF; placental growth factor; PLGF;
KM migration inhibitory factor; MIG; human; amphiregulin.
XX
OS Homo sapiens.
XX
PN WO2003097854-A2.
XX
PD 27-NOV-2003.
XX

PF 19-MAY-2003; 2003WO-US015711.
XX
PR 17-MAY-2002; 2002US-0380872P.
PR 24-FEB-2003; 2003US-0448874P.
PR 24-FEB-2003; 2003US-0448922P.
XX
PA (SUGEN-) SUGEN INC.
PI Morimoto A, Deprino S, O'farrell A, Smolich BD, Manning WC;
PI Walter SA, Schilling JW, Cherrington J;
XX
DR WPI: 2004-042604/04.
XX
PT Determining whether a test compound inhibits tyrosine kinase activity in
PT a mammal by exposing the mammal to the test compound and measuring in the
PT mammal the level of at least one of the measured proteins or mRNA
transcripts.
PS Disclosure; SEQ ID NO 51; 408bp; English.
XX
CC The invention relates to a novel method for determining whether a test
CC compound inhibits tyrosine kinase activity in a mammal comprising
CC measuring in the mammal the level of at least one of the proteins and/or
CC mRNA transcripts or genes for such proteins comprising type 1 plasminogen
CC activator inhibitor (PAI-1), TIMP-1 (tissue inhibitor of
CC metalloproteinase 1), vinculin, vascular endothelial growth factor
CC (VEGF), placental growth factor (PLGF), VEGF/PLGF heterodimers or
CC migration inhibitory factor (MIG), exposing the mammal to the test
CC compound and then measuring in the mammal the level of at least one of
CC the proteins and/or mRNA transcripts previously measured. The method of
CC the invention may be useful for determining whether a test compound
CC inhibits tyrosine kinase activity in a mammal. The current sequence is
CC that of the tyrosine kinase activity inhibition-related protein of the
invention.
XX
SQ Sequence 252 AA:
Query Match 100.0%; Score 1308; DB 8; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.2e-113;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRAPLPPAPVYVSLILGSGHYAAGLDNDYSGKREPFGSDHSDGFVTSRSESSG 60
DB 1 MRAPLPPAPVYVSLILGSGHYAAGLDNDYSGKREPFGSDHSDGFVTSRSESSG 60
QY 61 SEISPVSEMPSSSPSGADYDSEEDNEPQIPGYIVDDSVRVEQVVKPQNTSESENT 120
DB 61 SEISPVSEMPSSSPSGADYDSEEDNEPQIPGYIVDDSVRVEQVVKPQNTSESENT 120
QY 121 SDRPKRRKKGGKGNKRRNRKKNPCNAEFONFCIHGCKYIHLAVTCCKQOEYFGER 180
DB 121 SDRPKRRKKGGKGNKRRNRKKNPCNAEFONFCIHGCKYIHLAVTCCKQOEYFGER 180
QY 181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVITVAVITVQLRQYVRKTEGAEERKK 240
DB 181 CGEKSMTKTHSMIDSLSKIALAIAAFMSAVITVAVITVQLRQYVRKTEGAEERKK 240
QY 241 LROENGNVHAIA 252
DB 241 LROENGNVHAIA 252
RESULT 5
ADL12853
ID ADL12853 standard; protein; 252 AA.
XX
AC ADL12853;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human steroid-induced C3A liver cell protein #92.
XX
KM hepatotropic; Gene therapy; Wilson disease; liver disorder;
XX

KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
 XX Homo sapiens.
 OS
 XX US6673549-B1.
 PN
 XX 06-JAN-2004.
 PD
 XX
 PF 12-OCT-2001; 2001US-00976594.
 XX
 PR 12-OCT-2000; 2000US-0240409P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Furness LM, Buchbinder JL;
 XX WPI; 2004-068610/07.
 DR
 XX
 PT Combination useful for preparing a composition for treating liver
 PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
 PT comprising cDNAs that are differentially expressed in response to steroid
 XX treatment.
 PS Disclosure; SEQ ID NO 582; 141bp; English.
 CC The invention relates to a combination comprising cDNAs that are
 CC differentially expressed in response to steroid treatment. Also included
 CC are the following: a high throughput method for using a cDNA to detect
 CC differential expression of nucleic acids in a sample; and a high
 CC throughput method of screening molecules or compounds to allow
 CC ligand that specifically binds a cDNA. The sample is from a subject with
 CC Wilson disease and comparison of a standard defines a stage of that
 CC disease. The high throughput method of screening molecules or compounds
 CC to identify a ligand that specifically binds a cDNA comprises: combining
 CC the combination with molecules or compounds under conditions to allow
 CC specific binding; and detecting specific binding between each cDNA and at
 CC least one molecule or compound. The molecules or compounds are regulatory
 CC proteins. The combination is useful for preparing a composition for
 CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
 CC or hepatitis. The present sequence represents a human protein which is
 CC differentially expressed in steroid-induced C3A liver cells. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 XX at Seqdata.uspto.gov/sequence.html.
 SQ Sequence 252 AA;
 Query Match
 Best Local Similarity 100.0%; Score 1308; DB 8; Length 252;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRAPLPPAPVYLSILILSGHYAAGLDNDYSGRREPSGDSADGFEVTSRSMSSG 60
 DB 1 MRAPLPPAPVYLSILILSGHYAAGLDNDYSGRREPSGDSADGFEVTSRSMSSG 60
 QY 61 SEISPSVSEMPSSSEPSGADYDYSERYDNEPOIPGYVDSRYHGVVKKPQNTKESNT 120
 DB 61 SEISPSVSEMPSSSEPSGADYDYSERYDNEPOIPGYVDSRYHGVVKKPQNTKESNT 120
 QY 121 SIKPRKKKKGGNGNRRNRKKKPCNPFONFCIHCKCKIEHLAVTKCKCOQTFGER 180
 DB 121 SIKPRKKKKGGNGNRRNRKKKPCNPFONFCIHCKCKIEHLAVTKCKCOQTFGER 180
 QY 181 CGEKSMTKTHMIDSLKIALAIAAFMSAVITLAVAVITVQARKQYVKRYGEAEERKK 240
 DB 181 CGEKSMTKTHMIDSLKIALAIAAFMSAVITLAVAVITVQARKQYVKRYGEAEERKK 240
 QY 241 LRQENGVAHAIA 252
 DB 241 LRQENGVAHAIA 252
 QY 241 LRQENGVAHAIA 252
 DB 241 LRQENGVAHAIA 252

RESULT 6

ADJ75567
 ID ADJ75567 standard; protein; 252 AA.
 XX
 AC ADJ75567;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Marker gene related amino acid sequence SEQ ID NO:819.
 KW bronchial asthma; chronic obstructive pulmonary disease;
 KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 XX gene therapy; marker.
 OS Homo sapiens.
 XX
 PN EP1394274-A2.
 XX
 PD 03-MAR-2004.
 XX
 PF 04-AUG-2003; 2003EP-00254857.
 XX
 PR 06-AUG-2002; 2002JP-00229312.
 XX
 PR 20-MAR-2003; 2003JP-00077212.
 XX
 PA (GENO-) GENOX RES INC.
 XX
 PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
 XX WPI; 2004-193155/19.
 DR
 XX
 PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 XX healthy subject.
 XX
 PS Example 11; SEQ ID NO 819; 241bp; English.
 XX
 CC The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC stimulated with interleukin-13. Also described: (1) a reagent (1) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease;
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma or chronic
 CC obstructive pulmonary disease; (5) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for chronic
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC expression of the gene through an RNAi effect or an antibody recognizing
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilized to assay a marker gene. (1) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The method
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 252 AA;
 Query Match
 Best Local Similarity 100.0%; Score 1308; DB 8; Length 252;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAPLPAPVVLSTLITSGHYAAGLDLNDYSGKKEPSPGSDHSDGFEVTSRSEMSG 60
 DB 1 MRAPLPAPVVLSTLITSGHYAAGLDLNDYSGKKEPSPGSDHSDGFEVTSRSEMSG 60
 QY 61 SEISPVSEMPSSSPSSGADYDSEEDNEPQIPGYIVDVSVRVEQVVKPQNTSEST 120
 DB 61 SEISPVSEMPSSSPSSGADYDSEEDNEPQIPGYIVDVSVRVEQVVKPQNTSEST 120
 QY 121 SDKPKRKKKGGKNGKRNRRKKKPCNAEFQNCIHGECKYIHLAVTCKCOOEYFGER 180
 DB 121 SDKPKRKKKGGKNGKRNRRKKKPCNAEFQNCIHGECKYIHLAVTCKCOOEYFGER 180
 QY 181 CDKSKMTSHMIDSSLSKILAAIAAFMSAVITLAVAVITVQLRQVRYKTEGAEERKK 240
 DB 181 CGEKSMKTHMIDSSLSKILAAIAAFMSAVITLAVAVITVQLRQVRYKTEGAEERKK 240
 QY 241 LRQENGVAHAIA 252
 DB 241 LRQENGVAHAIA 252

RESULT 7
 ADP07715 ID ADP07715 standard; protein; 252 AA.
 XX
 AC ADP07715;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Human secreted protein, seq id 198.
 XX
 KW Cytostatic; antidiabetic; anorectic; gynaecological; antipsoriatic;
 KW dermatological; antihypertensive; antiepileptic; neuroprotective;
 KW neurologic; antiparkinsonian; nephrotoxic; human; secreted protein;
 KW diagnostic; pharmaceutical; cancer; lung; oesophageal; liver; diabetes;
 KW obesity; metabolic disorder; cardiovascular disorders; reproductive
 KW disorder; psoriasis; eczema; bronchitis; cystic fibrosis;
 KW atherosclerosis; benign prostatic hyperplasia; asthma;
 KW Alzheimer's disease; Parkinson's disease; renal disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO2004042000-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 16-MAY-2003; 2003WO-US015439.
 XX
 PR 17-MAY-2002; 2002US-0381592P.
 PR 12-JUN-2002; 2002US-038843P.
 PR 08-AUG-2002; 2002US-0401757P.
 PR 12-AUG-2002; 2002US-0402585P.
 PR 13-AUG-2002; 2002US-0402799P.
 PR 22-AUG-2002; 2002US-0404959P.
 PR 04-OCT-2002; 2002US-0415902P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Olsen H, Baker KP, Fiascella M, Wei P,
 PI Birse CE, Komatsoulis G, Choi GH, Moore PA, Gupta R, Shi Y,
 XX
 DR WPI; 2004-400658/37.
 DR N-PSDB; ADP07533.
 XX
 PT New human secreted polypeptides and nucleic acid molecules for
 PT diagnosing, preventing or treating disorders associated with the secreted
 PT proteins, such as cancer, diabetes, obesity, cardiovascular disorders or
 PT renal disorders.
 XX
 PS Claim 1; SEQ ID NO 198; 1157bp; English.
 XX
 CC The invention relates to a human secreted polypeptide for diagnosing,
 CC preventing or treating disorders associated with the secreted proteins.

CC The polypeptides and nucleic acid molecules of the invention are useful
 CC for preparing a diagnostic or pharmaceutical composition for diagnosing
 CC or treating a medical condition. These may be used for diagnosing,
 CC preventing or treating disorders related to the human secreted proteins,
 CC such as cancer (e.g. lung, oesophageal or liver cancer), diabetes,
 CC obesity, metabolic disorders, cardiovascular disorders, reproductive
 CC disorders, psoriasis, eczema, bronchitis, cystic fibrosis,
 CC atherosclerosis, benign prostatic hyperplasia, asthma, Alzheimer's
 CC disease, Parkinson's disease or renal disorders. Sequences given in
 CC records for ADP07710-ADP07891 represent human secreted proteins of the
 CC invention.
 XX
 SO Sequence 252 AA;
 Query Match 100.0%; Score 1308; DB 8; Length 252;
 Best Local Similarity 100.0%; Pred. No. 2,2e-113;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRAPLPAPVVLSTLITSGHYAAGLDLNDYSGKKEPSPGSDHSDGFEVTSRSEMSG 60
 DB 1 MRAPLPAPVVLSTLITSGHYAAGLDLNDYSGKKEPSPGSDHSDGFEVTSRSEMSG 60
 QY 61 SEISPVSEMPSSSPSSGADYDSEEDNEPQIPGYIVDVSVRVEQVVKPQNTSEST 120
 DB 61 SEISPVSEMPSSSPSSGADYDSEEDNEPQIPGYIVDVSVRVEQVVKPQNTSEST 120
 QY 121 SDKPKRKKKGGKNGKRNRRKKKPCNAEFQNCIHGECKYIHLAVTCKCOOEYFGER 180
 DB 121 SDKPKRKKKGGKNGKRNRRKKKPCNAEFQNCIHGECKYIHLAVTCKCOOEYFGER 180
 QY 181 CGEKSMKTHMIDSSLSKILAAIAAFMSAVITLAVAVITVQLRQVRYKTEGAEERKK 240
 DB 181 CGEKSMKTHMIDSSLSKILAAIAAFMSAVITLAVAVITVQLRQVRYKTEGAEERKK 240
 QY 241 LRQENGVAHAIA 252
 DB 241 LRQENGVAHAIA 252

RESULT 8
 ADR73587 ID ADR73587 standard; protein; 252 AA.
 XX
 AC ADR73587;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Wild-type full-length human AR polypeptide, seq id 1.
 XX
 KW Antipsoriatic; cytostatic; gene therapy; amphiregulin; AR; heparin;
 KW PAR34; PAR80; HuPAR34; chimeric; humanised; antibody; cancer; psoriasis;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO2004068931-A2.
 XX
 PD 19-AUG-2004.
 XX
 PF 06-FEB-2004; 2004WO-US004176.
 XX
 PR 07-FEB-2003; 2003US-0445640P.
 PR 30-DEC-2003; 2003US-0533901P.
 XX
 PA (PROT-) PROTEIN DESIGN LABS INC.
 XX
 PI Landolfi NF, Tsurushita N, Hinton PR, Kumar S;
 XX
 DR WPI; 2004-604310/58.
 XX
 PT New antibody that competitively inhibits binding of an amphiregulin (AR)
 PT polypeptide, useful for diagnosing or treating patients having cancer and
 PT psoriasis.

XX Example 1; SEQ ID NO 1; 99p; English.
 PS
 XX

CC The invention relates to an antibody that competitively inhibits binding
 CC of an amphitropin (AR) polypeptide to an antibody. Ar is a heparin-
 CC binding glycoprotein. The antibody comprises a heavy chain variable
 CC region having amino acid sequence of at least 60% identity to a sequence
 CC selected from ADR73588, ADR73590 and ADR73598, and a light chain variable
 CC region having amino acid sequence of at least 60% identity to a sequence
 CC selected from ADR73589, ADR73591 and ADR73600. The antibody binds to a sequence
 CC same AR epitope as that bound by an antibody selected from PAR34, PAR80,
 CC or HUPAR34, and is chimeric or humanised. The antibodies are useful for
 CC diagnosing or treating patients having cancer and psoriasis. The current
 CC sequence represents the wild-type full-length human AR polypeptide.
 XX
 SQ Sequence 252 AA;

Query Match
 Best Local Similarity 100.0%; Score 1308; DB 8; Length 252;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAPLPAPVYLILLIGSGHYAAGLDLNTYSGKREPSGDHSDGFEVTSRSEMSG 60
 DB 1 MRAPLPAPVYLILLIGSGHYAAGLDLNTYSGKREPSGDHSDGFEVTSRSEMSG 60
 QY 61 SEISPVSEMPSSSESGADYDSEEDNEPQIGYIVDSVREOVVKKPCKTESSENT 120
 DB 61 SEISPVSEMPSSSESGADYDSEEDNEPQIGYIVDSVREOVVKKPCKTESSENT 120
 QY 121 SDKPRKRRKGGKNGKRRNRKKNPCNAEFONFCIHGECKYIEHLAVTCKCOQEFGER 180
 DB 121 SDKPRKRRKGGKNGKRRNRKKNPCNAEFONFCIHGECKYIEHLAVTCKCOQEFGER 180
 QY 181 CGEKSMTKTHSMIDSSLSKIALAIAAFMSAVILTAVALTVQLRRQYVKGEGAEERKK 240
 DB 181 CGEKSMTKTHSMIDSSLSKIALAIAAFMSAVILTAVALTVQLRRQYVKGEGAEERKK 240
 QY 241 LRQENGVAHA 252
 DB 241 LRQENGVAHA 252

RESULT 9

ADR67228
 ID ADR67228 standard; protein; 252 AA.
 XX
 AC ADR67228;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human bladder cancer associated amino acid sequence.
 XX
 KW bladder cancer tissue; bladder cancer; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO2004076613-A2.
 XX
 PD 10-SEP-2004.
 XX
 PF 24-FEB-2004; 2004WO-DE000364.
 XX
 PR 26-FEB-2003; 2003DE-01009729.
 XX
 PA (HERR/) HERR A.
 PA (HINZ/) HINZMANN B.
 PA (DAHL/) DAHL E.
 PA (STAU/) STAU E.
 PA (PILARSKY C. (SPEC/)) PILARSKY C.
 PA (SPEC/)) SPECHT T.
 XX
 PI Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;

DR WPI; 2004-653385/63.
 XX
 XX

PT New nucleic acids, and encoded proteins, from bladder cancer tissue,
 PT useful for diagnosis, treatment and in screening for specific binding
 PT agents.
 XX
 XX

PS Claim 2; Fig 3; 112p; German.

CC The present invention describes nucleic acids (I) associated with bladder
 CC cancer tissue. Also described: (1) peptides and proteins (II) containing
 CC an amino acid sequence encoded by (I); (2) a method for diagnosing
 CC bladder cancer (BC) or monitoring its progression, that uses (I), (II)
 CC or agents (Z) that inhibit, or bind to, (I) or (II), labelled with a
 CC reporter; and (3) a method for treating BC that uses (I), (II) or (Z).
 CC (I) and (II) have cytostatic activity. (I) and (II) can be used to detect
 CC developing it; to screen for specific binding agents (Z), and to treat
 CC BC. (Z) are also useful as diagnostic and therapeutic agents. The present
 CC sequence represents a human amino acid sequence associated with bladder
 CC cancer, which is used in the exemplification of the present invention.
 XX
 SQ Sequence 252 AA;

Query Match
 Best Local Similarity 100.0%; Score 1308; DB 8; Length 252;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAPLPAPVYLILLIGSGHYAAGLDLNTYSGKREPSGDHSDGFEVTSRSEMSG 60
 DB 1 MRAPLPAPVYLILLIGSGHYAAGLDLNTYSGKREPSGDHSDGFEVTSRSEMSG 60
 QY 61 SEISPVSEMPSSSESGADYDSEEDNEPQIGYIVDSVREOVVKKPCKTESSENT 120
 DB 61 SEISPVSEMPSSSESGADYDSEEDNEPQIGYIVDSVREOVVKKPCKTESSENT 120
 QY 121 SDKPRKRRKGGKNGKRRNRKKNPCNAEFONFCIHGECKYIEHLAVTCKCOQEFGER 180
 DB 121 SDKPRKRRKGGKNGKRRNRKKNPCNAEFONFCIHGECKYIEHLAVTCKCOQEFGER 180
 QY 181 CGEKSMTKTHSMIDSSLSKIALAIAAFMSAVILTAVALTVQLRRQYVKGEGAEERKK 240
 DB 181 CGEKSMTKTHSMIDSSLSKIALAIAAFMSAVILTAVALTVQLRRQYVKGEGAEERKK 240
 QY 241 LRQENGVAHA 252
 DB 241 LRQENGVAHA 252

RESULT 10

AAAG75109
 ID AAG75109 standard; protein; 326 AA.
 XX
 AC AAG75109;
 XX
 DT 03-SEP-2001 (first entry)
 XX
 DE Human colon cancer antigen protein SEQ ID NO:5873.
 XX
 KW Human, colon cancer; colon cancer antigen; diagnosis; detection;
 XX colorectal carcinoma; chromosome 4.
 XX
 OS Homo sapiens.
 XX
 PN WO200122920-A2.
 XX
 PD 05-APR-2001.
 XX
 PF 28-SEP-2000; 2000WO-US026524.
 XX
 PR 29-SEP-1999; 99US-0157137P.
 PR 03-NOV-1999; 99US-0163280P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX WPI; 2001-235357/24.
 DR N-PSDB; AAH34514.
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers.
 XX
 XX Claim 11, Page 7390-7392; 9803pp; English.
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
 CC proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene therapy
 CC and vaccine production. N and P may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate P expression. For
 CC example, N and P may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of P by expressing inactive proteins or to
 CC supplement the patients own production of P. Additionally, N may be used
 CC to produce the colon cancer-associated PS, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the proteins. N and P
 CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922
 XX
 XX Sequence 326 AA;
 SQ
 Query Match 100.0%; Score 1308; DB 4; Length 326;
 Best Local Similarity 100.0%; Pred. No. 3.1e-113;
 Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRAPLPPAPVVLSTLITSGHYAAGLDLNDTYSGRKPPSGDHSADGFEVTSSEMSG 60
 DB 75 MRAPLPPAPVVLSTLITSGHYAAGLDLNDTYSGRKPPSGDHSADGFEVTSSEMSG 134
 QY 61 SEISPVSEMPSSSPSSGADYDSEBYDNEPQIPGYIVDVSVAVEQVVKPPQNTSESENT 120
 DB 135 SEISPVSEMPSSSPSSGADYDSEBYDNEPQIPGYIVDVSVAVEQVVKPPQNTSESENT 194
 QY 121 SDPKRRKKKGKGNKRRNRKKKPCNAEPONFCIHGECKYIEHLEAVTCKCOOEYFGER 180
 DB 195 SDPKRRKKKGKGNKRRNRKKKPCNAEPONFCIHGECKYIEHLEAVTCKCOOEYFGER 254
 QY 181 CGEKSMKTHSMIDSLSKIALAIAAFMSAVITLAVAVITVQLRQVRYKTEGAEERKK 240
 DB 255 CGEKSMKTHSMIDSLSKIALAIAAFMSAVITLAVAVITVQLRQVRYKTEGAEERKK 314
 QY 241 LROENGVAHAIA 252
 DB 315 LROENGVAHAIA 326

XX Key Location/Qualifiers
 FH Protein 101..171
 FT /label= amphiregulin
 FT
 PN DE3902157-A.
 XX
 XX 27-JUL-1989.
 PD
 XX
 XX 25-JAN-1989; 89DB-03902157.
 PP
 XX
 XX 25-JAN-1988; 88US-00148327.
 PR 15-APR-1988; 88US-00181884.
 PR 17-JAN-1989; 89US-00297816.
 XX
 XX (ONCO) ONCOGEN.
 PA (BRIM) BRISTOL-MYERS CO.
 XX
 PI Shoyab M, McDonald VL, Plozman G, Bradley JG;
 XX WPI; 1989-221730/31.
 DR N-PSDB; AAN90373.
 DR
 XX
 PT New growth regulating proteins designated amphiregulin - inhibiting
 PT tumours but stimulating some normal cells, also new precursor protein,
 PT encoding nucleic acid sequences and antibodies.
 XX
 XX Disclosure; Fig 17; 78pp; German.
 XX
 CC This sequence represents the precursor form of amphiregulin (ARPP). The
 CC corresp. DNA sequence is spliced to remove introns. The sequence is also
 CC processed to produce mature amphiregulin (AR). The mature protein is
 CC expressed using recombinant DNA methods and can be glycosylated or non-
 CC glycosylated. Abs can be produced which recognise an epitope of AR. AR is
 CC bifunctional growth regulator which inhibits growth of some cancers but
 CC stimulates growth of some normal cells. It is useful in the treatment
 CC diagnosis and monitoring of neoplastic diseases, for treating wounds etc.
 CC See also AAP90448, AAP90449, AAP90450, AAP90451, AAP90452, AAP90453,
 CC AAP90454, AAP90372, AAP90374, AAP90375 and AAP90376. (Updated on 25-MAR-
 CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC (Updated on 27-AUG-2003 to correct OS field.)
 XX
 XX Sequence 252 AA;
 SQ
 Query Match 99.0%; Score 1295; DB 1; Length 252;
 Best Local Similarity 99.2%; Pred. No. 3.5e-112;
 Matches 250; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MRAPLPPAPVVLSTLITSGHYAAGLDLNDTYSGRKPPSGDHSADGFEVTSSEMSG 60
 DB 1 MRAPLPPAPVVLSTLITSGHYAAGLDLNDTYSGRKPPSGDHSADGFEVTSSEMSG 60
 QY 61 SEISPVSEMPSSSPSSGADYDSEBYDNEPQIPGYIVDVSVAVEQVVKPPQNTSESENT 120
 DB 61 SEISPVSEMPSSSPSSGADYDSEBYDNEPQIPGYIVDVSVAVEQVVKPPQNTSESENT 120
 QY 121 SDPKRRKKKGKGNKRRNRKKKPCNAEPONFCIHGECKYIEHLEAVTCKCOOEYFGER 180
 DB 121 SDPKRRKKKGKGNKRRNRKKKPCNAEPONFCIHGECKYIEHLEAVTCKCOOEYFGER 180
 QY 181 CGEKSMKTHSMIDSLSKIALAIAAFMSAVITLAVAVITVQLRQVRYKTEGAEERKK 240
 DB 181 CGEKSMKTHSMIDSLSKIALAIAAFMSAVITLAVAVITVQLRQVRYKTEGAEERKK 240
 QY 241 LROENGVAHAIA 252
 DB 241 LROENGVAHAIA 252

RESULT 11
 AAP95447
 ID AAP95447 standard; protein; 252 AA.
 XX
 AC AAP95447;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 10-NOV-1989 (first entry)
 XX
 DE Amphiregulin precursor protein.
 XX
 KW Amphiregulin; growth regulator; neoplastic disease; bone resorption;
 KW immune response; arachidonic acid cascade.
 XX
 OS Human MCF 7 breast carcinoma cells.

RESULT 12
 ADCT1538
 ID ADCT1538 standard; protein; 235 AA.

Query Match	93.7%	Score 1226	DB 7	Length 235
Best Local Similarity	100.0%	Pred. No. 8.8e-106		
Matches 235	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	15	LIILGSHYAGLIDNDYSGKREPEPSGDHSDGFEVTSRSMGSGSEISPVSEMPSSSE	74	
Db	1	LIILGSHYAGLIDNDYSGKREPEPSGDHSDGFEVTSRSMGSGSEISPVSEMPSSSE	74	
QY	75	PSSGADYDYSEYNEPQIPGYITVDDSVRVQVKKPPONTKTESENTSPKPRKKKGGKNG	60	
Db	61	PSSGADYDYSEYNEPQIPGYITVDDSVRVQVKKPPONTKTESENTSPKPRKKKGGKNG	134	
QY	135	KNNRRKKKPNCAAEFONFCIHGECKYIEHLEAVTCKCOQEPGERGCKSMKTHSMIDS	120	
Db	121	KNNRRKKKPNCAAEFONFCIHGECKYIEHLEAVTCKCOQEPGERGCKSMKTHSMIDS	194	
QY	195	SLSKIALAIAAFMSAVIITAVVITVQARQVRYRYEGEAEERKKLRQENGWVH	249	
Db	181	SLSKIALAIAAFMSAVIITAVVITVQARQVRYRYEGEAEERKKLRQENGWVH	235	
RESULT 13				
ADN33949				
ID	ADN33949	standard, protein, 235 AA.		
XX	ADN33949			
AC	ADN33949			
XX	17-JUN-2004	(first entry)		
DT				
XX				
DE	Human novel protein NOV6a.			
XX				
KW	Human; NOV6; pancreatic polypeptide; angiotensin; interleukin-1;			
KW	endothelin-2; endocytosis; amphiregulin; metalloproteinase; insulin-like			
KW	metabolic disorder; diabetes; obesity; infectious disease; anorexia;			
KW	cancer; cancer-associated cachexia; neurodegenerative disorder;			
KW	Alzheimer's disease; Parkinson's disease; immune disorder;			
KW	hematologic disorder; dyslipidemia; metabolic disturbance;			
KW	metabolic syndrome X; wasting disorder; antibacterial agent; SNP;			
XX	single nucleotide polymorphism.			
OS	Homo sapiens.			
XX				
Key	Location/Qualifiers			
FT	Misc-difference 86	/note= "May be Tyr as the result of a single nucleotide		
FT	Misc-difference 89	polymorphism"		
FT	Misc-difference 89	/note= "May be Gly as the result of a single nucleotide		
FT	Misc-difference 124	polymorphism"		
FT	Misc-difference 124	/note= "May be Gly as the result of a single nucleotide		
FT	Misc-difference 124	polymorphism"		
XX	US2004038230-A1.			
XX				
PD	26-FEB-2004.			
XX				
PF	04-NOV-2002; 2002US-00287190.			
XX				
PR	05-NOV-2001; 2001US-0338626P.			
PR	28-NOV-2001; 2001US-0093601S.			
PR	28-NOV-2001; 2001US-033912P.			
PR	29-NOV-2001; 2001US-033912P.			
PR	03-DEC-2001; 2001US-0338100P.			
PR	19-APR-2002; 2002US-0338196P.			
PR	16-MAY-2002; 2002US-037806P.			
PR	07-AUG-2002; 2002US-0381043P.			
XX				
XX				
PA	(ALSO/) ANDERSON J P.			
PA	(ANDE/) ANDERSON D W.			

PA (CHAP/) CHAPOVAL A.
 PA (CHAU/) CHAUDHURI A.
 PA (EDIN/) EDINGER S R.
 PA (EISE/) EISEN A.
 PA (GANG/) GANGOLLI E A.
 PA (GERL/) GERLACH V.
 PA (GUOX/) GUO X S.
 PA (JITW/) JI W.
 PA (KRAA/) KRANTSOV N V.
 PA (LEIT/) LEITE M W.
 PA (LILL/) LI L.
 PA (MEZE/) MEZES P S.
 PA (MILL/) MILLET I.
 PA (OOIC/) OOI C E.
 PA (ORTT/) ORT T.
 PA (PADI/) PADIGARU M.
 PA (PATU/) PATTURAJAN M.
 PA (PENA/) PENNA C E A.
 PA (RAST/) RASTELLI L.
 PA (RIEG/) RIEGER D K.
 PA (SENG/) SENGGER K E Q.
 PA (SMIT/) SMITHSON G.
 PA (SPAD/) SPADERNA S K.
 PA (SPYT/) SPYTEK K A.
 PA (STON/) STONE D J.
 PA (TWOM/) TWOMLOW N.
 PA (VERN/) VERNET C A M.
 PA (VOSS/) VOSS E Z.
 PA (ZERR/) ZERNHUSEN B D.
 PA (ZHON/) ZHONG M.
 XX
 PI Alsobrook J P, Anderson D W, Boldog F L, Burgees C E, Chant J S;
 PI Chapoval A, Chaudhuri A, Edinger S R, Eisen A, Gangolli E A;
 PI Gerlach V, Guo X S, Ji W, Krantsov N V, Leite M W, Li L, Mezes P S;
 PI Millet I, Ooi C E, Ort T, Padigaru M, Paturajan M, Penna C E A;
 PI Rastelli L, Rieger D K, Sengger K E Q, Smithson G, Spaderna S K;
 PI Spytex K A, Stone D J, Twomlow N, Vernet C A M, Voss E Z, Zernhsen B D;
 PI Zhong M;
 XX
 DR WPI; 2004-213932/20.
 DR N-PSDB; ADN33948.
 XX
 PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,
 PT diagnosing and creating e.g. cancer, diabetes and Alzheimer's disease.
 XX
 PS Claim 2; SEQ ID NO 166; 129pp; English.
 XX
 PS The invention relates to 87 isolated NOVX polypeptides (NOV1-NOV5), their
 CC mature forms and their encoding polynucleotides having sequence
 CC similarity to pancreatic polypeptide (NOV1), angiotensin (NOV2),
 CC interleukin-1 (NOV3), endothelin-2 (NOV4), endozepine (NOV5),
 CC amphiregulin (NOV6) and metalloproteinase (NOV7). Also included the
 CC use of NOVX in the manufacture of a medicament for treating a human
 CC disease associated with NOVX, detecting NOVX in a sample via an
 CC immunoassay, identifying an agent that binds to the NOVX polypeptide,
 CC modulating the activity of NOVX, a vector comprising NOVX polynucleotide,
 CC a cell comprising the vector, an antibody that immunospecifically binds
 CC to NOVX, detecting the polynucleotide in a sample using a hybridisation
 CC assay, and producing the polypeptide comprising culturing the cell under
 CC conditions that lead to expression of the polypeptide. NOVX and its
 CC polynucleotide are used to prevent, diagnose or treat a medical condition
 CC in human related to the aberrant expression and activity of NOVX.
 CC polypeptides e.g. metabolic disorders, diabetes, obesity, infectious
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
 CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic
 CC disturbances associated with obesity, the metabolic syndrome X and
 CC wasting disorders associated with chronic diseases and various cancers.
 CC They may also be used as antibacterial agents. NOVX may also be used as
 CC an antigen in the production of antibodies and in assays to identify
 CC modulators (agonists and antagonists) of the expression and activity of
 CC NOVX. The present sequence represents a NOVX protein.

SO Sequence 235 AA:
 Query Match 93.7%; Score 1226; DB 8; Length 235;
 Best Local Similarity 100.0%; Pred. No. 8.8e-106;
 Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 LLLIGSGHYAAGLDINDTYSGKREPPSGDSADGFEVTSRSEMSGSEISPVSEMPSSSE 74
 DB 1 LLLIGSGHYAAGLDINDTYSGKREPPSGDSADGFEVTSRSEMSGSEISPVSEMPSSSE 60
 QY 75 PSSGADYDYESEYDNEPQIDGYIVDSVRYEQVVKPPQNTSESENTSDKRRKKKGKNG 134
 DB 61 PSSGADYDYESEYDNEPQIDGYIVDSVRYEQVVKPPQNTSESENTSDKRRKKKGKNG 120
 QY 135 KNRNRKKKPPCNAEFQNCIHGCKYIHLLEAVTCKCOOYFGERGCEGSMKTHSMIDS 194
 DB 121 KNRNRKKKPPCNAEFQNCIHGCKYIHLLEAVTCKCOOYFGERGCEGSMKTHSMIDS 180
 QY 195 SLSTKIALAATAAFMSAVITLTAVAVITVQLRRQYRKYEGABERKKLROENGVAH 249
 DB 181 SLSTKIALAATAAFMSAVITLTAVAVITVQLRRQYRKYEGABERKKLROENGVAH 235
 RESULT 14
 ID AAW73157 standard; protein; 226 AA.
 XX AAW73157;
 AC AAW73157;
 XX 20-JAN-1999 (first entry)
 DT 20-JAN-1999 (first entry)
 XX
 DE Human amphiregulin transmembrane glycoprotein.
 XX
 KE Amphiregulin; AR protein; epithelial cell growth; tumour cell growth;
 KW psoriasis; tumour marker; cancer diagnosis; therapy; human.
 XX
 OS Homo sapiens.
 XX
 XX US5830995-A.
 FN 03-NOV-1998.
 PD 03-NOV-1998.
 XX
 PE 18-MAY-1992; 92US-00885089.
 XX
 PR 25-JAN-1988; 88US-00148327.
 PR 15-APR-1988; 88US-00181884.
 PR 17-JAN-1989; 89US-00297816.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Plowman G D, Bradley J G, Shoyab M, McDonald V L;
 XX WPI; 1998-609313/51.
 DR WPI; 1998-609313/51.
 XX
 PT Amphiregulin proteins - useful for stimulating e.g. epithelial cell
 PT growth.
 XX
 PS Claim 11; Col 71-72; 76pp; English.
 XX
 CC This sequence represents the human amphiregulin (AR) transmembrane
 CC glycoprotein of the invention. The invention also relates to rat, mouse
 CC and bovine AR proteins. The AR proteins can be used to stimulate
 CC epithelial cell growth. They may be of therapeutic use in controlling
 CC tumour cell growth and in disorders such as psoriasis, and can be used as
 CC a tumour marker for cancer diagnosis
 XX
 SO Sequence 226 AA:
 Query Match 90.0%; Score 1177; DB 2; Length 226;
 Best Local Similarity 100.0%; Pred. No. 3.1e-101;
 Matches 226; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 LDINDTYSGKREPPSGDHSADGFEVTSRSEMSGSEISPVSEMPSSSEPSGADYDSE 86

```

Db      1 LDLDVTSKNEPFGSHSADGFEVTSRSESSGSEISFVSEMPSSSESGADYDSEE 60
Qy      87 YDNEPQIPGYIVDDSVREVEQVVKPPONKTESENTSPKPKKKKKGNKRRNRKKKQPC 146
Db      61 YDNEPQIPGYIVDDSVREVEQVVKPPONKTESENTSPKPKKKKKGNKRRNRKKKQPC 120
Qy      147 NAEFONFCIHGECKYIEHLFAVTCCKCOQYFGERCGEKMKTSHMIDSLSKTALAATA 206
Db      121 NAEFONFCIHGECKYIEHLFAVTCCKCOQYFGERCGEKMKTSHMIDSLSKTALAATA 180
Qy      207 FMSAVILTAVALITVQLRRQYVKYEGEAERKKLROENGNYHATA 252
Db      181 FMSAVILTAVALITVQLRRQYVKYEGEAERKKLROENGNYHATA 226

```

RESULT 15

```

AAW73160
ID      AAW73160 standard; protein; 247 AA.
XX
AC      AAW73160;
XX
DT      20-JAN-1999 (first entry)
XX
DE      Bovine amphiregulin.
XX
KW      Amphiregulin; AR protein; epithelial cell growth; tumour cell growth;
XX      psoriasis; tumour marker; cancer diagnosis; therapy; cow.
OS      Bos sp.
XX
PN      US5830995-A.
XX
PD      03-NOV-1998.
XX
PF      18-MAY-1992; 92US-00885089.
XX
PR      25-JAN-1988; 88US-00148327.
XX      15-APR-1988; 88US-00181884.
XX      17-JAN-1989; 89US-00297816.
XX
PA      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI      Plowman GD, Bradley JG, Shoyab M, McDonald VL;
XX
DR      WPI; 1998-609313/51.
XX
DR      N-PDDB; AAV59098.
XX
PT      Amphiregulin proteins - useful for stimulating e.g. epithelial cell
XX      growth.
XX
PS      Example; Fig 1a; 76pp; English.
XX
CC      This sequence represents the bovine amphiregulin (AR) of the invention.
XX      The invention also relates to rat, mouse and human AR proteins. The AR
XX      proteins can be used to stimulate epithelial cell growth. The AR
XX      therapeutic use in controlling tumour cell growth and in disorders such
XX      as psoriasis, and can be used as a tumour marker for cancer diagnosis
XX
SQ      Sequence 247 AA;

```

Query Match

Best local similarity 76.8%; Score 1004.5; DB 2; Length 247;
Matches 196; Conservative 19; Mismatches 31; Indels 7; Gaps 2;

```

Qy      1 MRAPLLPAPVVLSTLITSGHYAGLDINDTYSGKEPFGSDHSGADGFEVTSRSESSG 60
Db      1 MRAPLLPAPVVLSTLITSGHYAGLDINDTYSGKEPFGSDHSGADGFEVTSRSESSG 60
Qy      61 SEISPVSEMPSSSESGADYDSEYNEPQIPGYIVDDSVREVEQVVKPPONKTESENT 120
Db      61 SEISPVSEMPSSSESGADYDSEYNEPQIPGYIVDDSVREVEQVVKPPONKTESENT 114

```

```

Qy      121 SDRKRRKKKKGNKRRNRKKNPCNAEFONFCIHGECKYIEHLFAVTCCKCOQYFGER 180
Db      115 SDRKRRKKKKGNKRRNRKKNPCNAEFONFCIHGECKYIEHLFAVTCCKCOQYFGER 174
Qy      181 CGEKMKTSHMIDSLSKTALAATAFMSAVILTAVALITVQLRRQYVKYEGEAERK 239
Db      175 CGEKMKTSHMIDSLSKTALAATAFMSAVILTAVALITVQLRRQYVKYEGEAERK 234
Qy      240 KLRQENGNYHATA 252
Db      235 KLRQENGNYHATA 247

```

Search completed: February 26, 2005, 16:41:14
Job time : 167 secs


```

Db      121 SDPERRKKKKGGKNGKRRNRKKNPCNAEFQNCIHGECKYTEHLEAVTCKCOQDYFERR 180
QY      181 CGEKSMTKTHSMIDSSLSKIALAAIAAFMSAVILTAVAVITVOLRRQYVRKTEGEAEERKK 240
Db      181 CGEKSMTKTHSMIDSSLSKIALAAIAAFMSAVILTAVAVITVOLRRQYVRKTEGEAEERKK 240
QY      241 LROENGNAVHATA 252
Db      241 LROENGNAVHATA 252

```

RESULT 2

```

US-09-919-039-228
; Sequence 228, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: 2002-09-09
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 228
; LENGTH: 252
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 3478236CD1
US-09-919-039-228

```

```

Query Match      100.0%; Score 1308; DB 10; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.9e-110;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MRAPLPPAPVYVSLILIGSGHYAAGLDINDTYSGRREPFSGHSDGFEVTSRSEMSG 60
Db      1 MRAPLPPAPVYVSLILIGSGHYAAGLDINDTYSGRREPFSGHSDGFEVTSRSEMSG 60
QY      61 SEISPVSEMPSSSEPSGADYDYSEYDNEPQIPGYIVDDSVRVQVVKPONTTESSENT 120
Db      61 SEISPVSEMPSSSEPSGADYDYSEYDNEPQIPGYIVDDSVRVQVVKPONTTESSENT 120
QY      121 SDPERRKKKKGGKNGKRRNRKKNPCNAEFQNCIHGECKYTEHLEAVTCKCOQDYFERR 180
Db      121 SDPERRKKKKGGKNGKRRNRKKNPCNAEFQNCIHGECKYTEHLEAVTCKCOQDYFERR 180
QY      181 CGEKSMTKTHSMIDSSLSKIALAAIAAFMSAVILTAVAVITVOLRRQYVRKTEGEAEERKK 240
Db      181 CGEKSMTKTHSMIDSSLSKIALAAIAAFMSAVILTAVAVITVOLRRQYVRKTEGEAEERKK 240
QY      241 LROENGNAVHATA 252
Db      241 LROENGNAVHATA 252

```

RESULT 3

```

US-10-177-293-13
; Sequence 13, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillic, James
; APPLICANT: Glatz, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Ganuavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen

```

```

; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hottobagyl, Gabriel N.
; APPLICANT: Puozzati, Lajos
; APPLICANT: Sahin, Aysegül
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: PREVENTION, AND THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/177,293
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 252
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-177-293-13

```

US-10-177-293-13

```

Query Match      100.0%; Score 1308; DB 14; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.9e-110;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MRAPLPPAPVYVSLILIGSGHYAAGLDINDTYSGRREPFSGHSDGFEVTSRSEMSG 60
Db      1 MRAPLPPAPVYVSLILIGSGHYAAGLDINDTYSGRREPFSGHSDGFEVTSRSEMSG 60
QY      61 SEISPVSEMPSSSEPSGADYDYSEYDNEPQIPGYIVDDSVRVQVVKPONTTESSENT 120
Db      61 SEISPVSEMPSSSEPSGADYDYSEYDNEPQIPGYIVDDSVRVQVVKPONTTESSENT 120
QY      121 SDPERRKKKKGGKNGKRRNRKKNPCNAEFQNCIHGECKYTEHLEAVTCKCOQDYFERR 180
Db      121 SDPERRKKKKGGKNGKRRNRKKNPCNAEFQNCIHGECKYTEHLEAVTCKCOQDYFERR 180
QY      181 CGEKSMTKTHSMIDSSLSKIALAAIAAFMSAVILTAVAVITVOLRRQYVRKTEGEAEERKK 240
Db      181 CGEKSMTKTHSMIDSSLSKIALAAIAAFMSAVILTAVAVITVOLRRQYVRKTEGEAEERKK 240
QY      241 LROENGNAVHATA 252
Db      241 LROENGNAVHATA 252

```

RESULT 4

```

US-10-301-822-8
; Sequence 8, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.

```

```

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029P22RM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-8

Query Match      100.0%; Score 1308; DB 14; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.9e-110;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAPLLPPAPVYVLLILSGHYAAGLDNDYSGKREPSGDSADGFEVTSRSEMSG 60
    |||
Db 1 MRAPLLPPAPVYVLLILSGHYAAGLDNDYSGKREPSGDSADGFEVTSRSEMSG 60
    |||

QY 61 SEISPVSEMPSSSEPSGADYDYSEEDNEPQIPGYIVDSVRYVEQVVKPQNTSENT 120
    |||
Db 61 SEISPVSEMPSSSEPSGADYDYSEEDNEPQIPGYIVDSVRYVEQVVKPQNTSENT 120
    |||

QY 121 SDPKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||
Db 121 SDPKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||

QY 121 SDRKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||
Db 121 SDRKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||

QY 181 CGEKSMTKTHSMIDSLSKIALAAIAAFMSAVITAVAVITVQLRQYVRKYEGBAERKK 240
    |||
Db 181 CGEKSMTKTHSMIDSLSKIALAAIAAFMSAVITAVAVITVQLRQYVRKYEGBAERKK 240
    |||

QY 241 LROENGNVHAIA 252
    |||
Db 241 LROENGNVHAIA 252
    |||

RESULT 5
US-10-440-464-51
; Sequence 51, Application US/10440464
; Publication No. US20040018528A1
; GENERAL INFORMATION:
; APPLICANT: DEPRIMO, SAMUEL
; APPLICANT: O'FARRELL, ANNE-MARIE
; APPLICANT: MORIMOTO, ALYSSA
; APPLICANT: SMOLICH, BEVERLY
; APPLICANT: MANNING, WILLIAM
; APPLICANT: WALTER, SARAH
; APPLICANT: CHERINGTON, JULIE
; APPLICANT: SCHILLING, JIM
; TITLE OF INVENTION: NOVEL BIOMARKERS OF TYROSINE KINASE INHIBITOR EXPOSURE
; TITLE OF INVENTION: AND ACTIVITY IN MAMMALS
; FILE REFERENCE: 038602/1592
; CURRENT APPLICATION NUMBER: US/10/440,464
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: 60/380,872
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/448,922
; PRIOR FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: 60/448,874
; PRIOR FILING DATE: 2003-02-24
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 252
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-440-464-51

Query Match      100.0%; Score 1308; DB 15; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.9e-110;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAPLLPPAPVYVLLILSGHYAAGLDNDYSGKREPSGDSADGFEVTSRSEMSG 60
    |||
Db 1 MRAPLLPPAPVYVLLILSGHYAAGLDNDYSGKREPSGDSADGFEVTSRSEMSG 60
    |||

QY 61 SEISPVSEMPSSSEPSGADYDYSEEDNEPQIPGYIVDSVRYVEQVVKPQNTSENT 120
    |||
Db 61 SEISPVSEMPSSSEPSGADYDYSEEDNEPQIPGYIVDSVRYVEQVVKPQNTSENT 120
    |||

QY 121 SDPKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||
Db 121 SDPKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||

QY 121 SDRKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||
Db 121 SDRKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||

QY 181 CGEKSMTKTHSMIDSLSKIALAAIAAFMSAVITAVAVITVQLRQYVRKYEGBAERKK 240
    |||
Db 181 CGEKSMTKTHSMIDSLSKIALAAIAAFMSAVITAVAVITVQLRQYVRKYEGBAERKK 240
    |||

QY 241 LROENGNVHAIA 252
    |||
Db 241 LROENGNVHAIA 252
    |||

RESULT 6
US-10-106-698-5883
; Sequence 5883, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5883
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5883

Query Match      100.0%; Score 1308; DB 14; Length 326;
Best Local Similarity 100.0%; Pred. No. 4.1e-110;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAPLLPPAPVYVLLILSGHYAAGLDNDYSGKREPSGDSADGFEVTSRSEMSG 60
    |||
Db 75 MRAPLLPPAPVYVLLILSGHYAAGLDNDYSGKREPSGDSADGFEVTSRSEMSG 134
    |||

QY 61 SEISPVSEMPSSSEPSGADYDYSEEDNEPQIPGYIVDSVRYVEQVVKPQNTSENT 120
    |||
Db 135 SEISPVSEMPSSSEPSGADYDYSEEDNEPQIPGYIVDSVRYVEQVVKPQNTSENT 194
    |||

QY 121 SDPKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 180
    |||
Db 121 SDPKRRKKKGKGNKRRNKKKPCNAEPONFCIHGECKYIHLAIVTCCKQOEYFGER 254
    |||

QY 181 CGEKSMTKTHSMIDSLSKIALAAIAAFMSAVITAVAVITVQLRQYVRKYEGBAERKK 240
    |||
Db 255 CGEKSMTKTHSMIDSLSKIALAAIAAFMSAVITAVAVITVQLRQYVRKYEGBAERKK 314
    |||
```

QY 241 LROENGVAHA 252
DB 315 LROENGVAHA 326

RESULT 7

US-10-287-190-166
Sequence 166, Application US/10287190
Publication No. US20040038230A1
GENERAL INFORMATION:
APPLICANT: Alsobrook II, John P. et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-7808
CURRENT APPLICATION NUMBER: US/10/287,190
PRIOR FILING DATE: 2003-04-02
PRIOR APPLICATION NUMBER: 09/996,015
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/338626
PRIOR FILING DATE: 2001-11-05
PRIOR APPLICATION NUMBER: 60/373806
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/338196
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/333912
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/381043
PRIOR FILING DATE: 2002-05-16
PRIOR APPLICATION NUMBER: 60/401593
PRIOR FILING DATE: 2002-08-07
PRIOR APPLICATION NUMBER: 60/334300
NUMBER OF SEQ ID NOS: 194
SOFTWARE: Cnaseqblast version 0.1
SEQ ID NO 166
LENGTH: 235
TYPE: PRT
ORGANISM: Homo sapiens
US-10-287-190-166

Query Match

Best Local Similarity 93.7%; Score 1226; DB 15; Length 235;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 LLLIGSGHYAAGLDLNDYSGKREPRSSGSHSADGFEVTSRSMSGSEISPVSEMPSSSE 74
DB 1 LLLIGSGHYAAGLDLNDYSGKREPRSSGSHSADGFEVTSRSMSGSEISPVSEMPSSSE 74
QY 75 PSSGADYDSEEDNEPQIPGYIVDSVREGVVAP POKKTESSENTSKPRKKGKNG 60
DB 61 PSSGADYDSEEDNEPQIPGYIVDSVREGVVAP POKKTESSENTSKPRKKGKNG 134
QY 135 KRRNRKRRKKNPCNAEFONFCIHGECKYIEHLAVTCKCOQDFEGERGCKSMKTHSWIDS 120
DB 121 KRRNRKRRKKNPCNAEFONFCIHGECKYIEHLAVTCKCOQDFEGERGCKSMKTHSWIDS 194
QY 195 SLSKIAIAIAFMSAVILTAVAITVQLRQYVRKTEGSAERKKQLRONGVH 249
DB 181 SLKIAIAIAFMSAVILTAVAITVQLRQYVRKTEGSAERKKQLRONGVH 235

RESULT 8

US-10-022-609-15
Sequence 15, Application US/10022609
Publication No. US20030023035A1
GENERAL INFORMATION:
APPLICANT: Vandlen, Richard L.
TITLE OF INVENTION: Structure, Production and Use of
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 460 Point San Bruno Blvd

CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/022,609
FILING DATE: 17-Dec-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/440,401
FILING DATE: 12-MAY-1995
APPLICATION NUMBER: 08/330161
FILING DATE: 25-OCT-1994
APPLICATION NUMBER: 08/035430
FILING DATE: 22-MAR-1993
APPLICATION NUMBER: 07/705256
FILING DATE: 24-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 712C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 87 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-022-609-15

Query Match

Best Local Similarity 34.7%; Score 454; DB 14; Length 87;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 KKNPCNAEFONFCIHGECKYIEHLAVTCKCOQDFEGERGCKSMKTHSWIDSLSKIA 200
DB 1 KKNPCNAEFONFCIHGECKYIEHLAVTCKCOQDFEGERGCKSMKTHSWIDSLSKIA 60
QY 201 LAIAAFMSAVILTAVAITVQLRQY 227
DB 61 LAIAAFMSAVILTAVAITVQLRQY 87

RESULT 9

US-10-022-609-16
Sequence 16, Application US/10022609
Publication No. US20030023035A1
GENERAL INFORMATION:
APPLICANT: Vandlen, Richard L.
TITLE OF INVENTION: Structure, Production and Use of
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

```

Query Match 24.5%; Score 320; DB 14; Length 87;
Best Local Similarity 68.6%; Pred. No. 2,7e-21;
Matches 59; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

Cy 141 KKKPNCNAEPONFCIHGECKTIEHLEAVTCCQOEYFEERCGESKMTSHMIDSLSKIA 200
      KKKPNCNAEPONFCIHGECKTIEHLEAVTCCQOEYFEERCGESKMTSHMIDSLSKIA 60
Db 1 KKKPNCNAEPONFCIHGECKTIEHLEAVTCCQOEYFEERCGESKMTSHMIDSLSKIA 60

201 LAATAAFMSAVITLTAVAIVITQLRQ 226
      LAATAAFMSAVITLTAVAIVITQLRQ 86
Db 61 LAATIFVSNVSAIGITAVLARK 86

RESULT 10
US-10-201-945-11
; Sequence 11, Application US/10201945
; Publication No. US20020188110A1
GENERAL INFORMATION:
APPLICANT: Meisner, Paul S.
          Fuldner, Rebecca A.
          Adams, Mark D.
TITLE OF INVENTION: Transforming Growth Factor Alpha HI
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/201,945
FILING DATE: 25-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/471,377
FILING DATE: 23-Dec-1999
APPLICATION NUMBER: 08/915,096
FILING DATE: <Unknown>

```

RESULT 12
US-10-096-241-18
Sequence 18, Application US/10096241
Publication NO. US2002012759A1
GENERAL INFORMATION:
APPLICANT: Geating, David P.
Busfield, Samantha J.
TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
AND USES THEREFOR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston

STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/096,241
FILING DATE: 12-Mar-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/699,591
FILING DATE: 19-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07334/022001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-096-241-18

Query Match
Best Local Similarity 100.0%; Pred. No. 6.2e-17;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 KKNPCNAEFONFCIHGECKYIEHLAVYCKCOQYFGERCGEKM 186
Db 1 KKNPCNAEFONFCIHGECKYIEHLAVYCKCOQYFGERCGEKM 46

RESULT 13
US-09-817-647-10
Sequence 10, Application US/09817647
Patent No. US2002008229A1
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/817,647
FILING DATE: 26-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/107,979
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487

REFERENCE/DOCKET NUMBER: P1084R1-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: hAR.esf
LOCATION: 1-45
IDENTIFICATION METHOD:
OTHER INFORMATION:
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-817-647-10

Query Match
Best Local Similarity 20.1%; Score 263; DB 9; Length 45;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 KKNPCNAEFONFCIHGECKYIEHLAVYCKCOQYFGERCGEKM 186
Db 1 KKNPCNAEFONFCIHGECKYIEHLAVYCKCOQYFGERCGEKM 45

RESULT 14
US-09-877-665-10
Sequence 10, Application US/09877665
Patent No. US20020164680A1
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/877,665
FILING DATE: 08-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/109,206
FILING DATE: 30-Jun-1998
ATTORNEY/AGENT INFORMATION:
NAME: Conley, Delidre L.
REGISTRATION NUMBER: 36,487
REFERENCE/DOCKET NUMBER: P1084R1-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-2066
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
FEATURE:
NAME/KEY: hAR.esf
LOCATION: 1-45
IDENTIFICATION METHOD:
OTHER INFORMATION:
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-877-665-10

Query Match 20.1%; Score 263; DB 9; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.7e-16;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 KKNPCNAEFQNCIHGECCKYIEHLAVTCKCQCEYFGERCGEKSM 186
DB 1 KKNPCNAEFQNCIHGECCKYIEHLAVTCKCQCEYFGERCGEKSM 45

RESULT 15

US-10-136-573A-10
; Sequence 10, Application US/10136573A
; Publication No. US20020161200A1
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie Rose
; APPLICANT: Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and
; TITLE OF INVENTION: Uses therefor
; FILE REFERENCE: P1084R1C2
; CURRENT APPLICATION NUMBER: US/10/136, 573A
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 09/480,977
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: US 08/899,437
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: US 60/052,019
; PRIOR FILING DATE: 1997-07-09
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 10
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-136-573A-10

Query Match 20.1%; Score 263; DB 13; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.7e-16;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 KKNPCNAEFQNCIHGECCKYIEHLAVTCKCQCEYFGERCGEKSM 186
DB 1 KKNPCNAEFQNCIHGECCKYIEHLAVTCKCQCEYFGERCGEKSM 45

Search completed: February 26, 2005, 16:56:18
Job time : 136 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2005, 16:28:46 ; Search time 40 Seconds
(without alignments)
606.166 Million cell updates/sec

Title: US-10-774-076-1

Perfect score: 1308
Sequence: 1 MRAPLPPAPVVSLLILGS.....GEAEKKKLRQENGWHAIA 252

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	252	1	A34702
2	918	70.2	248	1	JH0612
3	874.5	66.9	243	2	S13296
4	211.5	16.2	208	1	A38432
5	211.5	16.2	208	1	A41914
6	185.5	14.2	208	1	UC1410
7	184.5	14.1	208	1	UC1409
8	162	12.4	84	2	S27162
9	144.5	11.0	380	2	G01639
10	143	10.9	346	2	T46914
11	137	10.5	141	2	S58117
12	119.5	9.1	177	2	A37408
13	117	8.9	860	2	UC5702
14	115	8.8	272	2	B71618
15	114.5	8.8	1207	1	EGHU
16	114	8.7	850	2	JC5700
17	113.5	8.7	162	2	S68401
18	112	8.6	1001	2	S30385
19	109	8.3	159	2	S27195
20	108	8.3	82	2	A61040
21	108	8.3	160	2	S39795
22	107	8.2	121	2	S58134
23	107	8.2	160	1	WFRU1
24	106	8.1	159	1	WFRU1
25	106	8.1	159	1	I57457
26	106	8.1	868	2	UC5701
27	104.5	8.0	178	2	JC1467
28	104	8.0	436	2	D88826
29	104	8.0	438	2	S28263

30	104	8.0	637	2	C43273	heregulin precursor
31	104	8.0	1670	2	T06754	DNA-directed RNA p
32	103.5	7.9	636	2	I61718	neu differentiatio
33	103.5	7.9	1133	1	EGRT	epidermal growth f
34	102.5	7.8	511	2	JC1404	CD81-box DNA-bind
35	102	7.8	661	2	T42754	hypothetical prote
36	101	7.7	639	2	I61719	neu differentiatio
37	100.5	7.7	296	2	A56943	sensory/motor neur
38	100.5	7.7	662	2	I61722	neu differentiatio
39	100	7.6	582	2	S69720	hypothetical prote
40	99	7.6	442	2	H86276	F14L17.4 protein -
41	99	7.6	645	2	B43275	heregulin, splice
42	98	7.5	140	2	T30766	growth factor - va
43	97.5	7.5	462	2	I38404	neu differentiatio
44	97.5	7.5	640	2	A43273	heregulin precursor
45	96.5	7.4	1217	1	EGMSMG	epidermal growth f

ALIGNMENTS

RESULT 1
A34702
amphiregulin precursor (validated) - human
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: A34702; B34702; A30057; A31230; A39788; A30056
R/Plowman, G.D.; Green, J.M.; McDonald, V.L.; Neubaer, M.G.; Diatche, C.M., To
Mol. Cell. Biol. 10, 1969-1981, 1990
A/Title: The amphiregulin gene encodes a novel epidermal growth factor-related p
A/Reference number: A34702; MUID:90220581; PMID:2325643
A/Accession: A34702
A/Molecule type: DNA
A/Residues: 1-252 <PI>
A/Cross-references: UNIPROT:P15514; GB:M30698
A/Accession: B34702
A/Molecule type: mRNA
A/Residues: 1-252 <PI2>
A/Cross-references: GB:M30704; NID:G179039; PIDN:AA51781.1; PID:G179040
R/Shoyab, M.; Plowman, G.D.; McDonald, V.L.; Bradley, J.G.; Todaro, G.J.
Science 243, 1074-1076, 1989
A/Title: Structure and function of human amphiregulin: a member of the epiderma
A/Reference number: A30057; MUID:89162036; PMID:2466334
A/Accession: A30057
A/Molecule type: protein
A/Residues: 101-112, 'D', 114-184 <SH1>
R/Shoyab, M.; McDonald, V.L.; Bradley, J.G.; Todaro, G.J.
Proc Natl. Acad. Sci. U.S.A. 85, 6528-6532, 1988
A/Title: Amphiregulin: a functional glycoprotein produced
A/Reference number: A94205; MUID:88320474; PMID:3413110
A/Accession: A31230
A/Molecule type: protein
A/Residues: 101-112, 'D', 114-131 <SH2>
R/Cook, P.W.; Mattox, P.A.; Keeble, W.W.; Pittelkow, M.R.; Plowman, G.D.; Sh
Mol. Cell. Biol. 11, 2547-2557, 1991
A/Title: A heparin sulfate-regulated human keratinocyte autocrine factor is
A/Reference number: A39788; MUID:91203875; PMID:2017164
A/Accession: A39788
A/Molecule type: protein
A/Residues: 101-136 <COO>
C/Comment: Amphiregulin is a secreted protein that is released from the tran
C/Genetics: A39788
A/Gene: GDB:AREG
A/Cross-references: GDB:119697; OMIM:104640
A/Map position: 4q13-q21
A/Introns: 21/1; 104/1; 171/1; 222/2
C/Superfamily: amphiregulin, EGF homology
C/Keywords: extracellular protein; glycoprotein; growth factor; transmembra
F/1-19/Domain: signal sequence #status predicted <Sig>
F/20-100/Domain: propeptide #status predicted <Pro>
F/101-184/Product: amphiregulin long form #status experimental <MA>
F/107-184/Product: amphiregulin short form #status experimental <MA>
F/146-181/Domain: EGF homology <EGF>

F:199-221/Domain: transmembrane #status predicted <TM>
F:222-252/Domain: intracellular #status predicted <INT>
F:30,113,119/Binding site: carbohydrate (asn) (covalent)
F:146-159,154-170,172-181/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 100.0%; Score 1308; DB 1; Length 252;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRAPLPPAPVYLSTLLIGSGHYAAGLDLNDITYSGRPPSGDHSADGFEVTSRSEMSG 60
1 MRAPLPPAPVYLSTLLIGSGHYAAGLDLNDITYSGRPPSGDHSADGFEVTSRSEMSG 60
SEISPVSEMPSSSESSGADYDYSEYDNEPQIPGYIVDVSRYVEQVVKPONTKESENT 120
SEISPVSEMPSSSESSGADYDYSEYDNEPQIPGYIVDVSRYVEQVVKPONTKESENT 120
SDPKPKKKKKGGKGNKRRKKKPCNAEFONFCIHGECKYIHLAVTCCKQOEYFGER 180
SDPKPKKKKKGGKGNKRRKKKPCNAEFONFCIHGECKYIHLAVTCCKQOEYFGER 180
CGEKMKTHTSMIDSLSKIALAAAFMSAVITLVAVITVQLRQYVRYKYEAEERKK 240
CGEKMKTHTSMIDSLSKIALAAAFMSAVITLVAVITVQLRQYVRYKYEAEERKK 240
LRQENGVAHAIA 252
LRQENGVAHAIA 252

RESULT 2

JH0612

amphiregulin precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C/Accession: JH0612; A56830

R/Source: H.; Yamaguchi, T.; Watanabe, S.

A/Title: Amphiregulin, Reg. Commun. 185, 103-109, 1992

A/Reference number: JH0612; MUID:92287078; PMID:1318038

A/Accession: JH0612

A/Molecule type: mRNA

A/Residues: 1-248 <SON>

A/Cross-references: UNIPROT:P31955; GB:D12648; DDBJ:D01182; NID:9220599; PIDN:BA02169.1

R/Date: Experimental source: SC2G cell

Mol. Endocrinol. 9, 691-705, 1995

A/Title: Amphiregulin is an implantation-specific and progesterone-regulated gene in the

A/Reference number: A56830; MUID:96174905; PMID:8592515

A/Accession: A56830

A/Molecule type: mRNA

A/Residues: 1-248 <DAS>

A/Note: authors translated the codon CCA for residue 84 as Glu

C/Comment: Amphiregulin is a secreted protein that is released from the transmembrane pr

C/Suprafamily: amphiregulin; EGF homology

F:1-19/Domain: extracellular protein; glycoprotein; growth factor; transmembrane pr

F:20-93/Domain: signal sequence #status predicted <PRO>

F:94-117/Product: amphiregulin long form #status predicted <SIG>

F:110-117/Product: amphiregulin short form #status predicted <MA>

F:192-217/Domain: EGF homology #status predicted <TM>

F:218-248/Domain: intracellular #status predicted <INT>

F:106/Binding site: carbohydrate (asn) (covalent) #status predicted

F:139-152,147-163,165-174/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 70.2%; Score 918; DB 1; Length 248;
Matches 177; Conservative 27; Mismatches 41; Indels 10; Gaps 3;

1 MRAPLPPAPVYLSTLLIGSGHYAAGLDLNDITYSGRPPSGDHSADGFEVTSRSEMSG 60

1 MRAPLPPAPVYLSTLLIGSGHYAAGLDLNDITYSGRPPSGDHSADGFEVTSRSEMSG 60
1 MRAPLPPAPVYLSTLLIGSGHYAAGLDLNDITYSGRPPSGDHSADGFEVTSRSEMSG 60
SEISPVSEMPSSSESSGADYDYSEYDNEPQIPGYIVDVSRYVEQVVKPONTKESENT 120
SEISPVSEMPSSSESSGADYDYSEYDNEPQIPGYIVDVSRYVEQVVKPONTKESENT 120
SDPKPKKKKKGGKGNKRRKKKPCNAEFONFCIHGECKYIHLAVTCCKQOEYFGER 180
SDPKPKKKKKGGKGNKRRKKKPCNAEFONFCIHGECKYIHLAVTCCKQOEYFGER 180
CGEKMKTHTSMIDSLSKIALAAAFMSAVITLVAVITVQLRQYVRYKYEAEERKK 240
CGEKMKTHTSMIDSLSKIALAAAFMSAVITLVAVITVQLRQYVRYKYEAEERKK 240
RRRLRQENGVAHAIA 248
RRRLRQENGVAHAIA 248

RESULT 3

S13296

Schwannoma-derived growth factor - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C/Accession: S13296

R/Kimura, H.; Fischer, W.H.; Schubert, D.

A/Title: Structure, expression and function of a schwannoma-derived growth facto

A/Reference number: S13296; MUID:91043095; PMID:2234093

A/Accession: S13296

A/Molecule type: mRNA

A/Residues: 1-243 <KIM>

A/Cross-references: UNIPROT:P24338; GB:X55183; NID:949595; PIDN:CA38967.1; PID:9

C/Suprafamily: amphiregulin; EGF homology

F:137-112/Domain: transmembrane protein

A/Title: Schwannoma-derived growth factor

A/Reference number: S13296; MUID:91043095; PMID:2234093

A/Accession: S13296

A/Molecule type: mRNA

A/Residues: 1-243 <KIM>

A/Cross-references: UNIPROT:P24338; GB:X55183; NID:949595; PIDN:CA38967.1; PID:9

C/Suprafamily: amphiregulin; EGF homology

F:137-112/Domain: transmembrane protein

A/Title: Schwannoma-derived growth factor

A/Reference number: S13296; MUID:91043095; PMID:2234093

A/Accession: S13296

A/Molecule type: mRNA

A/Residues: 1-243 <KIM>

A/Cross-references: UNIPROT:P24338; GB:X55183; NID:949595; PIDN:CA38967.1; PID:9

C/Suprafamily: amphiregulin; EGF homology

F:137-112/Domain: transmembrane protein

A/Title: Schwannoma-derived growth factor

A/Reference number: S13296; MUID:91043095; PMID:2234093

A/Accession: S13296

A/Molecule type: mRNA

A/Residues: 1-243 <KIM>

A/Cross-references: UNIPROT:P24338; GB:X55183; NID:949595; PIDN:CA38967.1; PID:9

C/Suprafamily: amphiregulin; EGF homology

F:137-112/Domain: transmembrane protein

A/Title: Schwannoma-derived growth factor

A/Reference number: S13296; MUID:91043095; PMID:2234093

A/Accession: S13296

A/Molecule type: mRNA

A/Residues: 1-243 <KIM>

A/Cross-references: UNIPROT:P24338; GB:X55183; NID:949595; PIDN:CA38967.1; PID:9

C/Suprafamily: amphiregulin; EGF homology

A:Accession: A38432
A:Molecule type: mRNA
A:Residues: 1-208 <HI>
A:Cross-references: UNIPROT:Q09075; GB:M60278; NID:g183866; PIDN:AAA35956.1; PID:g183867
R:HGhashiyama, S.; Lau, K.; Besner, G.E.; Abraham, J.A.; Klagesbrun, M.
J. Biol. Chem. 267, 6205-6212, 1992
A>Title: Structure of heparin-binding EGF-like growth factor. Multiple forms, primary str
A:Reference number: A37300; MUID:92210596; PMID:1556128
A:Accession: A37300
A:Molecule type: protein
A:Residues: 63-74, 'X', 76-84, 'X', 86-148 <HI>
C:Genetics:
A:Gene: GDB:DTR; DTS: HEGFL
A:Cross-references: GDB:119853; OMTM:126150
A:Map position: sq23-sq23
C:Superfamily: heparin-binding EGF-like growth factor; EGF homology
C:Keywords: heparin binding; transmembrane protein
F:108-143/Domain: EGF homology <EGF>

Query Match 16.2%; Score 211.5; DB 1; Length 208;
Best Local Similarity 36.7%; Pred. No. 2.3e-09;
Matches 55; Conservative 23; Mismatches 59; Indels 13; Gaps 4;

Oy 100 DSVREGVVKKPPONKTESENTSDKPKRKKGKGNRRNRKNPCNAEFGNCIHGEC 159
|::|||::|
Db DLRLTSSKKQALAT--PNKEHGKRKKKGKGG-----KKDPCLRKYKDCIHGEC 121
160 KYIEHLAVNTCKQQEYFGERCGEKSMKTHSIDSLKIALAAIAAFMSAVIITAVAVI 219
|::|||::|
Db KYVEELRAPSCICHGYGCHRGHGLSPVENRLTYDHITLIVAVAVLSSVCL---LVI 178
122
Oy 220 TVQLRRQYVRK-YEGEAERKCLRQENG 247
|::|||::|
Db VGLLMFRYHRRGYDVENEKEVKLGIMTSH 208
179

RESULT 5
A1914
diphtheria toxin receptor precursor - green monkey
C:Species: Cercopithecus aethiops (green monkey, grivet)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
A:Accession: A41914; JC4687
R:Naglich, J.G.; Methcarrall, J.E.; Russell, D.W.; Eidele, L.
Cell 69, 1051-1061, 1992
A>Title: Expression cloning of a diphtheria toxin receptor: identity with a heparin-bind
A:Reference number: A41914; MUID:92298386; PMID:1606612
A:Accession: A41914
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-208 <NAG>
A:Cross-references: UNIPROT:Q09118; GB:M93012
Biochem. Biophys. Res. Commun. 220, 675-680, 1996
R:Hooper, K.P.; Eidele, L.
A>Title: Glutamic acid 141 of the diphtheria toxin receptor (HB-EGF precursor) is critica
A:Reference number: JC4687; MUID:96183008; PMID:8607824
A:Accession: JC4687
A:Molecule type: mRNA
A:Residues: 1-208 <HO>
C:Comment: This precursor is a transmembrane glycoprotein which is cleaved at th e cell
C:Superfamily: heparin-binding EGF-like growth factor; EGF homology
C:Keywords: receptor; transmembrane protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:63-150/Product: diphtheria toxin receptor precursor #status predicted <MAT>
F:108-143/Domain: EGF homology <EGF>
F:160-184/Domain: transmembrane #status predicted <TM>

Query Match 16.2%; Score 211.5; DB 1; Length 208;
Best Local Similarity 36.7%; Pred. No. 2.3e-09;
Matches 55; Conservative 24; Mismatches 58; Indels 13; Gaps 4;

Oy 100 DSVREGVVKKPPONKTESENTSDKPKRKKGKGNRRNRKNPCNAEFGNCIHGEC 159
|::|||::|
Db DLRLTSSKKQALATPDK--BEHGKRKKKGKGG-----KKDPCLRKYKDCIHGEC 121
70

QY 160 KYIHLAATVCKCOQEFYGERCGEKSMTSMIDSSLSKIALAIAAFMSAVILITAVAVI 219
 Db 122 KYVELAAPSCTCHPGHGERCHGLSLPVENRLTYTHDTITLAVNAVVLSSVCL--LVI 178
 QY 220 TVQLRQYVRK--YEGEAERKKLROENGN 247
 Db 179 VGLIMFRYHRRGGYDVENEKVKLGMTNSH 208
 RESULT 6
 JCI1410
 heparin-binding EGF-like growth factor precursor - mouse
 C|Species: Mus musculus (house mouse)
 C|Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C|Accession: JCI1410; JCI4559
 R|Abraham, J.A.; Damm, D.; Bajardi, A.; Miller, J.; Klagebrun, M.; Ezekowitz, R.A.B.
 Biochem. Biophys. Res. Commun. 190, 125-133, 1993
 A|Title: Heparin-binding EGF-like growth factor: Characterization of rat and mouse CDNA
 A|Reference number: JCI409; MUID:93135756; PMID:7678488
 A|Accession: JCI1410
 A|Molecule type: mRNA
 A|Residues: 1-208 <AB>
 A|Cross-references: UNIPROT:Q06186; GB:L07264; NID:g192999; PIDN:AAA37542.1; PID:g19300
 R|Harding, P.A.; Briscock, D.R.; Shen, L.; Clissman-Combs, M.A.; Beener, G.E.
 Gene 169, 291-292, 1996
 A|Title: Characterization of the gene encoding murine heparin-binding epidermal growth
 A|Reference number: JCI459; MUID:96194822; PMID:8647467
 A|Accession: JCI459
 A|Molecule type: DNA
 A|Residues: 1-208 <HAR>
 A|Cross-references: GB:U39189
 C|Comment: This factor is a member of the epidermal growth factor family. Its binding a
 C|Genetics:
 A|Gene: mHB-EGF
 A|Introns: 16/1; 74/1; 133/2; 185/2
 C|Superfamily: heparin-binding EGF-like growth factor; EGF homology
 C|Keywords: growth factor; heparin binding; transmembrane protein
 F|1-23/Domain: signal sequence #status predicted <SIG>
 F|24-208/Product: heparin-binding EGF-like growth factor #status predicted <MAT>
 F|102-143/Domain: EGF homology <EGF>
 F|161-184/Domain: transmembrane #status predicted <TM>
 Query Match 14.2%; Score 185.5; DB 1; Length 208;
 Best Local Similarity 32.5%; Pred. No. 2.5d-07;
 Matches 39; Conservative 28; Mismatches 48; Indels 5; Gaps 2;
 QY 124 PRRKKKKGKNGKRNRRKKKNPCNAEPFQFCHGECKYIHLAATVCKCOQEFYGERCGE 183
 Db 86 PSKERNNKKKKKKGKGLGKRRDPCILKRYKDYCIHGECKRYLQFPFPPSCKCLPGYHGRCHG 145
 QY 184 KSMKTHMIDSSLSKIALAIAAFMSAVILITAVAVITVQLRQYVRK--YEGEAERKKL 241
 Db 146 LTLLEVENPLTYHTTVLAVNAVVLSSVCL--LVIVGLIMFRYHRRGGYDLSEBKKVL 202
 RESULT 7
 JCI1409
 heparin-binding EGF-like growth factor precursor - rat
 C|Species: Rattus norvegicus (Norway rat)
 C|Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C|Accession: JCI1409
 R|Abraham, J.A.; Damm, D.; Bajardi, A.; Miller, J.; Klagebrun, M.; Ezekowitz, R.A.B.
 Biochem. Biophys. Res. Commun. 190, 125-133, 1993
 A|Title: Heparin-binding EGF-like growth factor: Characterization of rat and mouse CDNA
 A|Reference number: JCI409; MUID:93135756; PMID:7678488
 A|Accession: JCI1409
 A|Molecule type: mRNA
 A|Residues: 1-208 <ABR>
 A|Cross-references: UNIPROT:Q06175; GB:L05489; NID:g204289; PIDN:AAA81780.1; PID:g20429
 C|Superfamily: heparin-binding EGF-like growth factor; EGF homology
 C|Keywords: growth factor; heparin binding; transmembrane protein
 F|1-23/Domain: signal sequence #status predicted <SIG>

```

A:Cross-references: GB:U39189
A:Comment: This factor is a member of the epidermal growth factor family. Its binding a
C:Genetics:
A:Gene: mHB-EGF
A:Introns: 16/1; 74/1; 133/2; 185/2
C:Superfamily: heparin-binding EGF-like growth factor; EGF homology
C:Keywords: growth factor; heparin binding; transmembrane protein
F.1-23/Domain: signal sequence #status predicted <SIG>
F.24-206/Product: heparin-binding EGF-like growth factor #status predicted <MKT>
F.108-143/Domain: EGF homology <EGP>
F.161-164/Domain: transmembrane #status predicted <TM>

Query Match      14.2%; Score 185.5; DB 1; Length 208;
      Best Local Similarity 32.5%; Pred. No. 2.5e-07;
      Matches 39; Conservative % 28; Mismatches 48; Indels 5; Gaps 2;

QY      124 PKKKKKGKNGKRRNRKKNPCAEFQNCIHCECKYIEHLNAVCKGQEXFGEECK 183
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      86 PSKERNKKKKKKGLGKKDKPCLRKYDKCIHCECKYLDSPRTSPCKCLPGHGHCHG 145

QY      184 KSMKTHSMIDSLSKIALAAIAFMFSAVITLAVAVITVQDRQYRK--YEGAEERKKL 241
      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      146 LTLVEVNPILTYDHTTVLAVAVAVLSVCL--LVIVGLMFRYRHRGGYDLSEBRKVL 202

RESULT 7
JC1409
heparin-binding EGF-like growth factor precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: JC1409
R:Abraham, J.A.; Damm, D.; Bajardi, A.; Miller, J.; Klagsbrun, M.; Ezekowitz, R.A.B.
Biochem. Biophys. Res. Commun. 190, 125-133, 1993
A:Title: Heparin-binding EGF-like growth factor: Characterization of rat and mouse CDNA
A:Reference number: JC1409; MUID:93135756; PMID:7678488
A:Accession: JC1409
A:Molecule type: mRNA
A:Residues: 1-208 <RAT>
A:Cross-references: UNIPROT:Q06175; GB:I05489; NID:G204289; PIDN:AAA81780.1; PID:G20429
C:Superfamily: heparin-binding EGF-like growth factor; EGF homology
C:Keywords: growth factor; heparin binding; transmembrane protein
F.1-23/Domain: signal sequence #status predicted <SIG>

```

F:24-208/Product: heparin-binding EGF-like growth factor #status predicted <MAT>
F:108-143/Domain: EGF homology <EGF>
F:161-184/Domain: transmembrane #status predicted <TM>

Query Match

Best Local Similarity 14.3%; Score 184.5; DB 1; Length 208;
Matches 40; Conservative 28; Mismatches 47; Indels 5; Gaps 2;

QY 124 PKRKKGGKNGKRRRRKKKPNCAFEONFCIHGECKYTEHLEAVTCQCOEYFGEKCE 183
DB 86 PGEKNGKKRKKKGGKRRKDPCKRYKQYCIHGECKRYKELRISCHCLPGYHQRCCHG 145
QY 184 KSNKTHSMIDSLSKIALAIAAFMSAVILTAVALTVQLRROYRK--YGEAEERKKL 241
DB 146 LTLPEVNPPLYTYDHTTVLAVAVVSSVCL---LVYGLMFRYHRGGYDLESEKVKL 202

RESULT 8

S27162
heparin-binding epidermal growth factor - pig (fragment)

C/Species: Sus scrofa domestica (domestic pig)

C/Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004

C/Accession: S27162; S25158

R/Vaughan, T.J.; Pascall, J.C.; Brown, K.D.

Biochem. J. 287, 681-684, 1992

A/Title: Tissue distribution of mRNA for heparin-binding epidermal growth factor.

A/Reference number: S27162; MUID:93075016; PMID:1445231

A/Accession: S27162

A/Molecule type: mRNA

A/Residues: 1-84 <VAN>

A/Cross-references: UNIPROT:00180; EMBL:X67295; NID:g1969; PIDN:CAA47709.1; PID:g1970

C/Superfamily: heparin-binding EGF-like growth factor; EGF homology

C/Keywords: growth factor; heparin binding

F:10-45/Domain: EGF homology <EGF>

Query Match

Best Local Similarity 12.4%; Score 162; DB 2; Length 84;
Matches 29; Conservative 17; Mismatches 30; Indels 0; Gaps 0;

QY 141 KKNPCNAEFONFCIHGECKYTEHLEAVTCQCOEYFGEKCEKMTSMIDSLSKIA 200
DB 5 KKRDPCKRYKQYCIHGECKRYKELRISCHCLPGYHQRCCHG 145
QY 201 LAATAAFMSAVILTAVALTVQLRROYRK--YGEAEERKKL 241
DB 65 LAVAVVSSVCLVLT 80

RESULT 9

G01639
transmembrane protein - human

C/Species: Homo sapiens (man)

C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004

C/Accession: G01639

R/Eid, D.W.

Submitted to the EMBL Data Library, January 1995

A/Reference number: G07997

A/Accession: G01639

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-380 <RIB>

A/Cross-references: UNIPROT:Q13086; EMBL:U19878; NID:g755465; PIDN:AAA64622.1; PID:g7554

F:275-310/Domain: EGF homology <EGF>

Query Match

Best Local Similarity 11.0%; Score 144.5; DB 2; Length 380;
Matches 54; Conservative 45; Mismatches 95; Indels 55; Gaps 9;

QY 24 AAGLDINDTYSGRRPFSGDHSADGFEVTSRESEM-----SS 59
DB 139 ARGPCVSDNGSGSGE---GEERGSABEYHRKSKGCPCKYARCEDEANVGCVCNIDCS 195
QY 60 GSEIPEVMSPPSSSGADYDYSEEDNEPQIFGYIVDS-VKVEQV-VKPPONKTES 117

DB 196 GSEFNPVADSGSS-----YNN---PCFVRASCTKQDIDIRLGHCTDT 238

QY 118 ENTS-----DKPRKKKGGKNGKRRNR---KKKPNCAFEONFCIHGECKYTEHLEAVT 169

DB 239 DDTSLGKDDGLQYRPDVPDASDQREDYVIGNHMPCEPUNNGCIHGKCEFYLLRRAS 238

QY 170 CKCOEYFGEKCEKMTSMIDSLSKIALAIAAFMSAVILTAVALTVQLRROYRK 229

DB 229 CRCESGTGGHC-EKTPSILYVVPSPKRLTHVLAIAIGAVALIATVATWCTTRKCPK 357

QY 230 KYEGEAER 238

DB 358 NNRGRQKQ 366

RESULT 10

T46914
hypothetical protein DKFZ564L1878.1 - human

C/Species: Homo sapiens (man)

C/Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004

C/Accession: T46914

R/Ottewaelder, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.

Submitted to the Protein Sequence Database, February 2000

A/Reference number: Z24136

A/Accession: T46914

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-346 <AAA>

A/Cross-references: UNIPROT:Q9NS55; EMBL:AL157430

A/Experimental source: fetal brain; clone DKFZ564L1878

C/Genetics:

A/Note: DKFZ564L1878.1

Query Match

Best Local Similarity 10.9%; Score 143; DB 2; Length 346;
Matches 47; Conservative 46; Mismatches 104; Indels 34; Gaps 8;

QY 15 LILILSGHYAGLDINDTYSGRRPFSGDHSADGFEVTSRESEM-----GSEIPEVSE 68
DB 127 ILVVEGSCA-----TDAGSG-----SGDGVHESGETSOKETSTCICPGACEDDAE 176

QY 69 -----MPSSEPSGADYDYSEEDNEPQI--PGYIVDSVAVQVVPKPKTE 116
DB 177 DWVCVNCIDSCQINNFPCASD---GKSYDAQCIRKASCCQKRIKIMSLGRCDWTTT 233

QY 117 SENTSDKPRKKKGGKNGKRRNRKTN--PCNAFEONFCIHGECKYTEHLEAVTCQCOQ 174
DB 234 TTSEEDGHVARTDYAENANKLEBSARBNHIPCPEHYNGFCMHGKCEHSTIMQEPSCKDA 293

QY 175 EYPERGCEKEMKTHSMIDSLSKIALAIAAFMSAVILTAVALTVQLR 225
DB 294 GYTQHGCEKQDYVLYVVPGV-RFOYVTLAAVIGTQIIVICVVLCTIR 343

RESULT 11

S58117
EGF-like growth factor, precursor - green monkey

C/Species: Cercopithecus aethiops (green monkey, grivet)

C/Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004

C/Accession: S58117

R/Loukianov, E.V.; Loukianov, T.I.; Wiedlocha, A.; Olenev, S.

Submitted to the EMBL Data Library, July 1995

A/Description: Short form of heparin binding EGF-like growth factor.

A/Reference number: S58117

A/Accession: S58117

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-141 <LOU>

A/Cross-references: UNIPROT:Q28218; EMBL:X89728; NID:g1052560; PIDN:CAA61880.1; PID:g105

C/Superfamily: heparin-binding EGF-like growth factor; EGF homology

Query Match

Best Local Similarity 10.5%; Score 137; DB 2; Length 141;

A:Residues: 1-1207 <BRL>
A:Cross-references: UNIPROT:P01133; EMBL:X04571; NID:G31120; PIDN:CAA28240.1; PID:G31121
A:Note: 708-Met was also found
A:Note: Intron positions were also determined
R:Gregory, H.; Preston, B.M.
Int. J. Pept. Protein Res. 9, 107-118, 1977
A:Title: The primary structure of human urogastrone
A:Reference number: A01388; MUID:77117897; PMID:300079
A:Accession: A01388
A:Molecule type: protein
A:Residues: 971-1023 <GRE>
A:Note: some of the molecules lack Arg-1023
R:Funya, M.; Akashi, S.; Hirayama, K.
Biochem. Biophys. Res. Commun. 163, 1100-1106, 1989
A:Title: The primary structure of human EGF produced by genetic engineering, studied by
A:Reference number: A33517; MUID:89391964; PMID:2789514
A:Accession: A33517
A:Molecule type: protein
A:Residues: 971-1023 <FUR>
R:Tsuikumo, K.; Nakamura, H.; Sakamoto, S.
Biochem. Biophys. Res. Commun. 145, 126-133, 1987
A:Title: Purification and characterization of high molecular weight human epidermal grow
A:Reference number: A29721; MUID:87241488; PMID:3297054
A:Accession: A29721
A:Molecule type: protein
A:Residues: 829-834, 'X', 836-839, 'X', 841-845, 'X', 847-848 <TSU>
A:Note: this is the amino-terminal sequence of a high molecular weight form of EGF, isol
R:Syoboda, M.; Baehofler, A.; Schwind, P.; Bade, E.; Rasched, I.; Przybylski, M.
Biochim. Biophys. Acta 1206, 33-41, 1994
A:Title: Structural characterization and biological activity of recombinant human epider
A:Reference number: S45282; MUID:9424778; PMID:8186248
A:Accession: S45282
A:Molecule type: protein
A:Residues: 'M', 971-1023 <SVO>
A:Note: expressed recombinant protein
A:Accession: S45283
A:Molecule type: protein
A:Residues: 'MKKYP', 970-1023 <SV2>
A:Note: expressed recombinant protein
C:Comment: Epidermal growth factor (EGF) stimulates the proliferation and differentiat
C:Comment: Epidermal growth factor (EGF) stimulates the proliferation and differentiat
C:Comment: EGF is released in the pancreas, small intestine, mammary gland, and (in some
C:Comment: The EGF precursor is found in kidney as a receptor-like membrane-bound protei
C:Genetics:
A:Gene: GDB:EGF
A:Cross-references: GDB:119105; OMIM:131530
A:Map position: 4q25-4q25
A:Intons: 43/1; 109/3; 170/2; 246/2; 314/1; 356/3; 397/1; 438/1; 525/3; 575/2; 6
C:Superfamily: epidermal growth factor precursor; EGF homology; LDL receptor WYTD-contai
C:Keywords: duplication; growth factor; tandem repeat; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-1207/Product: epidermal growth factor proporotein, membrane-bound form #status predi
F:23-1032/Domain: extracellular #status predicted <EXT>
F:43-479/Region: EGF precursor long repeat <LR1>
F:46-85/Domain: LDL receptor WYTD-containing repeat homology <YW01>
F:86-127/Domain: LDL receptor WYTD-containing repeat homology <YW02>
F:128-169/Domain: LDL receptor WYTD-containing repeat homology <YW03>
F:170-211/Domain: LDL receptor WYTD-containing repeat homology <YW04>
F:212-256/Domain: LDL receptor WYTD-containing repeat homology <YW05>
F:257-301/Domain: LDL receptor WYTD-containing repeat homology <YW06>
F:318-354/Domain: EGF homology <EG1>
F:360-395/Domain: EGF homology <EG2>
F:401-436/Domain: EGF homology <EG3>
F:439-476/Domain: EGF homology <EG4>
F:480-954/Region: EGF precursor long repeat <LR2>
F:483-523/Domain: LDL receptor WYTD-containing repeat homology <YW07>
F:524-566/Domain: LDL receptor WYTD-containing repeat homology <YW08>
F:567-609/Domain: LDL receptor WYTD-containing repeat homology <YW09>
F:610-653/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F:654-694/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F:695-737/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F:745-780/Domain: EGF homology <EG5>
F:835-868/Domain: EGF homology <EG6>

F:874-910/Domain: EGF homology <EG7>
F:916-951/Domain: EGF homology <EG8>
F:971-1023/Product: epidermal growth factor #status experimental <EGF>
F:976-1012/Domain: EGF homology <EG9>
F:1033-1057/Domain: transmembrane #status predicted <TM>
F:1058-1207/Domain: intracellular #status predicted <INT>
F:318-330,325-339,341-354,360-371,367-380,382-395,401-412,408-421,423-436,439-451,447-4
F:ide bonds: #status predicted
F:976-990,984-1001,1003-1012/Disulfide bonds: #status experimental

Query Match 8.8%; Score 114.5; DB 1; Length 1207;
Best Local Similarity 22.5%; Pred. No. 0.6;
Matches 50; Conservative 22; Mismatches 95; Indels 55; Gaps 9;

Db 1 MRAPLPAPVVLSTLILSGHYAAGLDINDTSGKREPFSGDHSADGPEVRSSEMSG 60
876 MGVPVCPAPS-----SKINTEGGVVC--RSEEGVQ-----DGHICDIDECQLG 919
QY 61 SEISPVSEMPSSSEPSGADYDYSEBYDNEPQIPGYVDDSVRVQVAKPPQNKTESENT 120
Db 920 --VHSGENASCTNTEGG-----YTCWCAGRLESPGLICPDST-----PPHLREDPH- 965
QY 121 SPKPRKKKKGGKNGKRRKKNPCNMFQNECTH-GECKYIEHLEAVTCKCOQYFGE 179
Db 966 -----HYVRNDSRCPUSHSGYCLHDGCMYIEALDKYACNCVGVYIGE 1010
QY 180 RGEKSMKTHSMIDSLKIALAIAAFMSAVILTRAVAVTV 221
Db 1011 RQGYRLKWEELRHAGHGQ-----QOKYIVAVCVVL 1043

Search completed: February 26, 2005, 16:45:00
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2005, 15:50:20 ; Search time 174 Seconds
(without alignments)
741.632 Million cell updates/sec

Title: US-10-774-076-1

Perfect score: 1308

Sequence: 1 MRAPLPPAPVVLSTLIGS.....GEAERKKLQENGNVHAA 252

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1308	100.0	252	1	AMPR_HUMAN
2	1059.5	81.0	247	2	Q9BDH4
3	941	71.9	216	2	Q9BDH3
4	918	70.2	248	1	AMPR_MOUSE
5	877	67.0	248	2	Q62561
6	874.5	66.9	243	1	AMPR_RAT
7	607.5	46.4	165	2	Q8K1E3
8	435.5	33.3	210	2	Q645M5
9	348	26.6	77	2	P79199
10	224.5	17.2	239	2	Q6N1Y9
11	211.5	16.2	208	1	HBBG_CERAE
12	211.5	16.2	208	1	HBBG_HUMAN
13	208.5	15.9	208	1	HBBG_MOUSE
14	198.5	15.2	212	2	Q9W7C5
15	195	14.9	132	2	Q9B6C2
16	187	14.3	208	1	Q9R0C7
17	185.5	14.2	208	1	HBBG_MOUSE
18	184.5	14.1	208	1	HBBG_RAT
19	183	14.0	105	2	Q99NM7
20	150	11.5	374	2	Q9QYM6
21	145	11.1	368	2	Q9P2Y9
22	145	11.1	374	2	Q8N2R5
23	145	11.1	374	2	Q9UIK5
24	144.5	11.0	380	2	Q130B6
25	143.5	11.0	341	2	Q8N3T8
26	143.5	11.0	380	2	Q8IYR6
27	143	10.9	346	2	Q9NS55
28	142	10.9	373	2	Q9QYV1
29	137	10.5	141	2	Q28218
30	136	10.4	354	2	Q9J5S1
31	134	10.2	247	2	Q8C536

Result 1	ID	AMPR_HUMAN	STANDARD;	PRT;	252 AA.
AC	P15514				
DT	01-APR-1990	(Rel. 14, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	05-JUL-2004	(Rel. 44, Last annotation update)			
DE	Amphiregulin precursor (AR) (Colorectal cell-derived growth factor) (CRDGF).				
GN	Name=AREG;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90220581; PubMed=2325643;				
RA	Plowman G.D., Green J.M., McDonald V.L., Neubauer M.G., Distèche C.M.,				
RA	Todaro G.J., Shoyab M.,				
RT	"The amphiregulin gene encodes a novel epidermal growth factor-related				
RL	protein with tumor-inhibitory activity.";				
RL	Mol. Cell. Biol. 10:1969-1981(1990).				
RN	[2]				
RP	SEQUENCE FROM N.A. AND VARIANTS VAL-80 AND CYS-81.				
RA	Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,				
RA	Myamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,				
RA	Schackwitz W.S., Sherwood J.K., Wittrik L.A., Nickerson D.A.,				
RT	"NIHES-SNPs, environmental genome project, NIHES ES15478, Department				
RT	of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";				
RL	Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shellen C.M., Schuler G.D.,				
RA	Altshuler S.F., Zeeberg B., Bueltow K.H., Schaefer C.P., Bhac N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schaefer T.E.,				
RA	Bromenstein M.J., Udell T.B., Tohbiyuki S., Carninci P., Prange C.,				
RA	Raba S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,				
RA	Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Klares S., Worley K.C., Hale S., Garcia A.M., Gay L.V., Hulyk S.W.,				
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Pahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,				
RA	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				

ALIGNMENTS

32	133	10.2	256	2	Q8RP7	Q8RP7 mus musculus
33	133	10.2	372	2	Q6PEP7	Q6PEP7 mus musculus
34	128	9.8	172	2	Q915B0	Q915B0 xenopus lae
35	124	9.5	272	2	Q8T7T2	Q8T7T2 plasmodium
36	124	9.5	272	2	Q8T7T3	Q8T7T3 plasmodium
37	124	9.5	272	2	Q8T7T8	Q8T7T8 plasmodium
38	123.5	9.4	177	2	Q9JDM4	Q9JDM4 ratius norv
39	123	9.4	272	2	Q8ITM3	Q8ITM3 plasmodium
40	120	9.2	272	2	Q810F0	Q810F0 plasmodium
41	120	9.2	272	2	Q810I8	Q810I8 plasmodium
42	120	9.2	272	2	Q8ITL6	Q8ITL6 plasmodium
43	120	9.2	272	2	Q8ITL6	Q8ITL6 plasmodium
44	120	9.2	272	2	Q70ID2	Q70ID2 plasmodium
45	120	9.2	272	2	Q70IE8	Q70IE8 plasmodium

[illegible]

SQ SEQUENCE 247 AA; 27336 MW; 6D1FA6BDFE4B9957.CRC64;
 Query Match 81.0%; Score 1059.5; DB 2; Length 247;
 Best Local Similarity 80.6%; Pred. No. 7.9e-74;
 Matches 204; Conservative 19; Mismatches 23; Indels 7; Gaps 2;

QY 1 MRAPLPAPVYVSLILISGHTAAGLDLNDTYSCKREPSGSHSADGFEVTSREMSG 60
 DB 1 MRAPLPAPVYVSLILISGHTAAGLDVNGTSSGKEPPSGDHGAFAFEVTSREMSG 60
 QY 61 SEISPVSEMPSSSESGADYDSEEDNEPQIPGYIVDSVVEQVVKPDKTESSENT 120
 DB 61 SEAPPASEMP-----SGSDYDAEEDNEPHISGYIVDSVVEQVVKPDKRTESSENT 114
 QY 121 SDPKRKKKKGGKGNKRRNKKKPCNAEFQNCIHGECKYIEHLAATCKCOQYFGER 180
 DB 115 SDPKRKKKKGGKGNKRRNKKKPCNAEFQNCIHGDCKYIEHLAATCKCYQDYFGER 174
 QY 181 CGEKSMTKTHMISDSLSKIALAIAAFMSAVILTAIV-VITVQLRQYVKRYGGEAEERK 239
 DB 175 CGEKSMTKTHMISDSLSKIALAIAAFVMSFTALAVITTYLRKKYFREYEGAEERK 234
 QY 240 KLRQENGNAIA 252
 DB 235 KLRQENANNAIA 247

RESULT 3
 Q9BDH3 PRELIMINARY; PRT; 216 AA.
 AC Q9BDH3;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Amphiregulin short form.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kim J.G., Vallet J.L., Rohrer G.A., Christenson R.K.;
 RC Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 DR EMBL: AY028311; AAK27729.1; -;
 DR HSSP: Q99075; 1XDT.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR006209; EGF_1like.
 DR InterPro: IPR006210; IEGF.
 DR Pfam: PF00008; EGF_1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS50026; EGF_3; 1.
 KM EGF-like domain.
 FT CHAIN 1..95 178 amphiregulin short form.
 SQ SEQUENCE 216 AA; 23571 MW; SDA65FPI87ED89DD.CRC64;

Query Match 71.9%; Score 941; DB 2; Length 216;
 Best Local Similarity 81.7%; Pred. No. 9.3e-65;
 Matches 179; Conservative 16; Mismatches 18; Indels 6; Gaps 1;

QY 1 MRAPLPAPVYVSLILISGHTAAGLDLNDTYSCKREPSGSHSADGFEVTSREMSG 60
 DB 1 MRAPLPAPVYVSLILISGHTAAGLDVNGTSSGKEPPSGDHGAFAFEVTSREMSG 60
 QY 61 SEISPVSEMPSSSESGADYDSEEDNEPQIPGYIVDSVVEQVVKPDKTESSENT 120
 DB 61 SEAPPASEMP-----SGSDYDAEEDNEPHISGYIVDSVVEQVVKPDKRTESSENT 114
 QY 121 SDPKRKKKKGGKGNKRRNKKKPCNAEFQNCIHGECKYIEHLAATCKCOQYFGER 180
 DB 115 SDPKRKKKKGGKGNKRRNKKKPCNAEFQNCIHGDCKYIEHLAATCKCYQDYFGER 174

QY 181 CGEKSMTKTHMISDSLSKIALAIAAFMSAVILTAIV 219
 DB 175 CGEKSMTKTHMISDSLSKIALAIAAFVMSFTALAV 213

RESULT 4
 ID AMPR_MOUSE STANDARD; PRT; 248 AA.
 AC P31955;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE Amphiregulin precursor (AR) (Schwannoma-derived growth factor) (SDGF).
 GN Name-Areg; Synonyms=8d9f;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92287078; PubMed=1318038;
 RA Sonoda H., Yamaguchi T., Watanabe S.;
 RT "Androgen-responsive expression and mitogenic activity of schwannoma-derived growth factor on an androgen-dependent Shionogi mouse mammary carcinoma cell line."
 RL Biochem. Biophys. Res. Commun. 185:103-109 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96174905; PubMed=8592515; DOI=10.1210/me.9.6.691;
 RA Das S.K., Chakraborty I., Paria B.C., Wang X.N., Ploegman G.D.,
 RA Dey S.K.;
 RT "Amphiregulin is an implantation-specific and progesterone-regulated gene in the mouse uterus."
 RL Mol. Endocrinol. 9:691-705 (1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cecum;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki I., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nikaide I., Osato N., Saito R., Suzuki H., Yamashita H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Balderrelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schriml L.M., Kandian A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Bradt D., Brasic V., Chochia C., Corbani L.B., Cousins S.,
 RA Dalia E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustinchik S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Kanagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglocz D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.U., Perle G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelin A., Schneider C., Semple C.A., Seton M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Walestec C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirokane-Kishikawa T., Kono H., Nakamura M., Sakazume N.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straube R.L., Feigold E.A., Grouse L.H., Dey J.G.,
 RA Kraemer R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Braletton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalon D.K., Muzny K.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Matra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Autocrine growth factor as well as a mitogen for a broad
 CC range of target cells, among which astrocytes, Schwann cells and
 CC fibroblasts.
 CC -1- INDUCTION: Androgen-dependent.
 CC -1- SIMILARITY: Belongs to the amphiregulin family.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D12648; BAB02169.1; -
 DR EMBL: L41357; AAB00472.1; -
 DR EMBL: AK018590; BAB31296.1; -
 DR EMBL: BC009138; AAB09138.1; -
 DR FIR: JH0612; JH0612.
 DR HSSP: Q99075; 1XDT.
 DR MGD: MGI:88068; Areg.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR006209; EGF_1like.
 DR PROSITE: PS00022; EGF_1; 1.
 DR PROSITE: PS01186; EGF_2; FALSE_NEG.
 DR PROSITE: PS50026; EGF_3; 1.
 DR CYTOKINE: EGF-like domain; Glycoprotein; Growth factor; Signal;
 KW Transmembrane.
 FT SIGNAL 1 26 Potential.
 FT PROPEP 27 99
 FT CHAIN 100 248 Amphiregulin.
 FT DOMAIN 135 175 EGF-like.
 FT TRANSMEM 192 215 Potential.
 FT DISULFID 139 152 By similarity.
 FT DISULFID 147 163 By similarity.
 FT DISULFID 174 174 By similarity.
 FT CARBOHYD 106 106 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 241 241 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 248 AA; 27549 MW; 98C61A1B0E75A64E CRC64;
 Query Match 70.2%; Score 918; DB 1; Length 248;
 Best Local Similarity 69.4%; Pred. No. 6-63;
 Matches 177; Conservative 27; Mismatches 41; Indels 10; Gaps 3;
 QY 1 MRAPLPPAVVLLIIGSGHYAAGLDINDYSGKREPFSGDHSADGFEVTSRSEMSG 60
 DB 1 MRPPLPLARVLLVLLVLSGSHYAAALINDPSGKSGSLSGDHSAGL-----ELSVG 54
 QY 61 SETSPVSESPSSSEPSGADYDYSEEDNEPQIPGYIVDDSVRYEYQVYKPKONTSESEN 120
 DB 55 REVSTISEMPSGSELSTG-DYDSEEDNEPQISGYITDDSVRYEYQVYKPKONTSEKKS 113
 QY 121 SDKPRKKGKNGKNGRRKRNKPCNAEFONFCIHGECKYIHLAEVTCCKOQYFGER 180
 DB 121 SDKPRKKGKNGKNGRRKRNKPCNAEFONFCIHGECKYIHLAEVTCCKOQYFGER 180

DB 114 TEKPRKKGKNGKNGRRKRNKPCNAEFONFCIHGECKYIHLAEVTCCKOQYFGER 173
 QY 181 CGEKSMTSMDSLSKIALAIAFMSAVILTA--VITVQLRQYRKYEAGEAE 237
 DB 174 CGEKSMTSEDDKLSKIAVAVAIVFSAIILAIIGIGIVITVLMRYRREYEGTEE 233
 QY 238 RKKLRQENGVAAIA 252
 DB 234 RKKLRQENGVAAIA 248
 RESULT 5
 ID 062561 PRELIMINARY; PRT; 248 AA.
 AC 062561;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Schwannoma derived growth factor.
 OS Mesocricetus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 CX NCBI_TaxID=10038;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97041064; PubMed=8975671;
 RA Gurskaya N.G., Shagin D.A., Luk'yanov K.A., Vagner L.L.,
 RA Shubman M.S., Muecklin E.A., Molnova E.V., Tatosian A.G.,
 RA Luk'yanov S.A., Sverdlov E.D.;
 RT "Cloning cDNA for the ha-SDF gene from a Syrian hamster cell line
 RT with increased metastatic potential using subtractive hybridization."
 RL Bioorg. Khim. 22:425-431(1996).
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 DR EMBL: Z66535; CA91439.1; -
 DR HSSP: Q99075; 1XDT.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR006209; EGF_1like.
 DR InterPro: IPR006210; IEGF.
 DR Pfam: PF00008; EGF_1.
 DR SMART: SM00181; EGF_1.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE: PS50026; EGF_3; 1.
 KW EGF-like domain
 SQ SEQUENCE 248 AA; 27381 MW; CD920DD40B3D1063 CRC64;
 Query Match 67.0%; Score 877; DB 2; Length 248;
 Best Local Similarity 69.4%; Pred. No. 9-6-60;
 Matches 177; Conservative 26; Mismatches 42; Indels 10; Gaps 5;
 QY 1 MRAPLPPA-PVVLISLIIIGSGHYAAGLDINDYSGKREPFSGDHSADGFEVTSRSEMSG 59
 DB 1 MRPPLPLARVLLVLLVLSGSHYAAALINDPSGKSGSLSGDHSAGL-----SV 54
 QY 60 GSEISPVSESPSSSEPSGADYDYSEEDNEPQIPGYIVDDSVRYEYQVYKPKONTSESEN 119
 DB 55 GREVSTISEMPSGSELSTG-DYDSEEDNEPQISGYITDDSVRYEYQVYKPKONTSESEN 113
 QY 120 TSDKPRKKGKNGKNGRRKRNKPCNAEFONFCIHGECKYIHLAEVTCCKOQYFGER 178
 DB 114 TSDKPRKKGKNGKNGRRKRNKPCNAEFONFCIHGECKYIHLAEVTCCKOQYFGER 173
 QY 179 ERGERSKMTSMDSLSKIALAIAFMSAVILTA--VITVQLRQYRKYEAGEAE 237
 DB 174 ERGERSKMTSMDSLSKIALAIAFMSAVILTA--VITVQLRQYRKYEAGEAE 233
 QY 238 RKKLRQENGVAAIA 252
 DB 234 RKKLRQENGVAAIA 248
 RESULT 6
 AMR_RAT

ID AMPR RAT STANDARD; PRT; 243 AA.
AC P24338;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 05-JUN-2004 (Rel. 44, Last annotation update)
GN Amphiregulin precursor (AR) (Schwannoma-derived growth factor) (SDGF).
OS Name=Arreg; Synonyms=Sdgf;
NCBI_TaxID=10116;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=91043095; PubMed=2234093; DOI=10.1038/348257a0;
RA Kimura H., Fischer W.H., Schubert D.;
RT "Structure, expression and function of a schwannoma-derived growth factor";
RL Nature 348:257-260(1990).
CC -1- FUNCTION: Autocrine growth factor as well as a mitogen for a broad range of target cells, among which astrocytes, Schwann cells and fibroblasts.
CC -1- SIMILARITY: Belongs to the amphiregulin family.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; X55183; CAA3867.1; -.
DR FTR; S13296; S13296.
DR HSP; Q99075; 1XDT.
DR RGD; 2149; Arreg.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR006209; EGF_Like.
DR Pfam; PF00008; EGF_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS0186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
KM Cytochrome; direct protein sequencing; EGF-like domain; Glycoprotein;
KM Growth factor; Signal; Transmembrane.
FT SIGNAL 1 24
FT PROPEP 25 96
FT CHAIN 97 243 Amphiregulin.
FT DOMAIN 133 173 EGF-like.
FT TRANSMEM 137 150 Potential.
FT DISULFID 145 161 By similarity.
FT DISULFID 163 172 By similarity.
FT CARBOHYD 30 30 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 103 103 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 236 236 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 243 AA; 26633 MW; F59A8173D4921559 CRC64;

Query Match 66.9%; Score 874.5; DB 1; Length 243;
Best Local Similarity 69.7%; Pred. No. 1.5e-59;
Matches 177; Conservative 24; Mismatches 40; Indels 13; Gaps 5;

QY 1 MRAPLPPAPVYSLILIGSGHYAAGLDNDYSGKREPFGSDHSAQFEVTSSEMSG 60
DB 1 MRPSTLSLAVSLVLSLIGSGHYAAGLDNDYSGKREPFGSDHSAQFEVTSSEMSG 51
QY SEISPVSEMPSSSPSGADYDYSEEDNEPQIPGYIVDSVREOVVKKPONTSESENT 120
DB SEISPVSEMPSSSPSGADYDYSEEDNEPQIPGYIVDSVREOVVKKPONTSESENT 110
QY 121 SDRKRRKKGGKGNKRRN-KKKNPQNAEFQNCIHGECKYIEHLAVTCKCOQEEYFE 179
DB 111 SEKPKRRKKGGKGNKRRN-KKKNPQNAEFQNCIHGECKYIEHLAVTCKCHQDYFE 170

QY 180 RCGEKMKTSHMIDSSLSKIALAIAFMSAVILTAVAVT-VOLRRQYRYKEGEAEER 238
DB 171 RCGEKMKTSHMIDSSLSKIALAIAFMSAVILTAVAVT-VOLRRQYRYKEGEAEER 229
QY 239 KKLQENGNVAAIA 252
DB 230 RRLQENGTAAIA 243

RESULT 7
Q8K3E3 PRELIMINARY; PRT; 165 AA.
ID Q8K3E3
AC Q8K3E3
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Amphiregulin (Fragment).
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang X., Reese J., Paria B.C.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY135447; AA095451.1; -.
DR HSP; Q99075; 1XDT.
FT NON_TER 1 1
FT NON_TER 165 165
SQ SEQUENCE 165 AA; 17986 MW; 7572D02022DAFB26 CRC64;

Query Match 46.4%; Score 607.5; DB 2; Length 165;
Best Local Similarity 70.6%; Pred. No. 3.6e-39;
Matches 120; Conservative 16; Mismatches 25; Indels 9; Gaps 4;

QY 5 LPPAPVYSLILIGSGHYAAGLDNDYSGKREPFGSDHSAQFEVTSSEMSGSEI 63
DB 1 LPPAPVYSLILIGSGHYAAGLDNDYSGKREPFGSDHSAQFEVTSSEMSGSEI 54
QY SPVSEMPSSSPSGADYDYSEEDNEPQIPGYIVDSVREOVVKKPONTSESENTSDK 123
DB STVSEMPSSSPSGADYDYSEEDNEPQIPGYIVDSVREOVVKKPONTSESENTSDK 113
QY 124 PKRRKKGGKGNKRRN-KKKNPQNAEFQNCIHGECKYIEHLAVTCKC 172
DB 114 PKRRKKGGKGNKRRN-KKKNPQNAEFQNCIHGECKYIEHLAVTCKC 163

RESULT 8
Q645M5 PRELIMINARY; PRT; 210 AA.
ID Q645M5
AC Q645M5
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Amphiregulin.
GN Name=AREG;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archaeoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Ovary;
RA Wang Y., Li J., Leung F.C.;
RT "Chicken Amphiregulin Gene: cDNA Cloning, Promoter Analysis, and Regulation of its mRNA Expression in the Chicken Ovary";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY25836; AA021467.1; -.
SQ SEQUENCE 210 AA; 23665 MW; B977D786740150CB CRC64;

Query Match 33.3%, Score 435.5; DB 2; Length 210;
 Best Local Similarity 48.6%; Pred. No. 9.1e-26;
 Matches 88; Conservative 25; Mismatches 59; Indels 9; Gaps 4;

QY 70 PSSSEPSGADYDYSEEDNEPQIPGIYDDSVREVEYVPPQNKTESNTSKPRKXK 129
 DB 35 PREPEPARGSDYE-EEFEYEAFLAHQVLVDLVRVEVYKPKAKGSEKNAKPPRRKN 93
 QY 130 GKKKGRNRNRKKNPCNAFONFCIHGECKYIEHLAVTCKQOEYFGERGCKSKMTH 189
 DB 94 KGR-----NKKGTPEMEYKNFCIHGECVYLQHMATCKCYQYFGERGCEQFMKTQ 147
 QY 190 SMID-SLSKIALAALAFMSAVILPAVAVITQLAROVYRKAGEAERKTLROENGAV 248
 DB 148 RKVDVADYSKTVLVVAVALLSSISFAVILITIVQVRKCKCPYE-BKEERKTLROENRNS 206
 QY 249 H 249
 DB 207 H 207

RESULT 9

PRELIMINARY; PRT; 77 AA.
 ID P79199
 AC P79199;
 BT 01-MAY-1997 (T-EMBLrel. 03, Created)
 DT 01-MAY-1997 (T-EMBLrel. 03, Last sequence update)
 DE 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 OS Amphiregulin (Fragment).
 OC Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Caprinae; Ovis.
 NCBI_Taxid=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Welsh Mountain; TISSUE=Mammary gland;
 RX MEDLINE=97179075; PubMed=9027362; DOI=10.1016/S0303-7207(96)03967-6;
 RA Foreyth I.A., Taylor J.A., Keable S., Hurvey A., Lennard S.;
 RT "Expression of amphiregulin in the sheep mammary gland.";
 RL Mol. Cell. Endocrinol. 126:41-48(1997).
 CC -1. SIMILARITY: Contains 1 EGF-like domain.
 DR EMBL; Y08830; CAA70973.1; -.
 DR HSSP; Q99075; 1XDT.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR006210; IEGF.
 DR Pfam; PF00008; EGF_1.
 DR SMART; SMO0181; EGF_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR EGF-like domain.
 KM EGF-like domain.
 FT NON TER 1 1
 FT NON TER 77 77
 SQ SEQUENCE 77 AA; 9060 MW; 00BE7B59AE685F25 CRC64;

Query Match 26.6%; Score 348; DB 2; Length 77;
 Best Local Similarity 77.9%; Pred. No. 1.5e-19;
 Matches 60; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 108 VKPPQNTSESENTSDPKRRKKGKGNKRRNRKKNPCNAFONFCIHGECKYIEHLA 167
 DB 1 VKPKKNTSEKTSKDPKRRKKGKGNKRRNRKKNLCTDFQNFCHGKCTFLEQLET 60
 QY 168 VTCKCOOEYFGERGCK 184
 DB 61 VSCOCYPEYFGERGCK 77

RESULT 10

PRELIMINARY; PRT; 239 AA.
 ID Q6NY79
 AC Q6NY79;

DT 05-JUL-2004 (T-EMBLrel. 27, Created)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
 DE LOC407664 protein (Fragment).
 GN Name=LOC407664;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_Taxid=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shennan C.M., Schlier G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Caavaant T.L., Scheetz T.E.,
 RA Brownstein M.J., Utsid T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lonnellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosek S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Young A.C., Shevchenko Y.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalski U., Smalins D.E., Schnerch A., Schein J.R.,
 RA Jones S.J., Mitra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC066705; AAH6705.1; -.
 DR InterPro; IPR001336; EGF_1.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_Like.
 DR Pfam; PF00008; EGF_1.
 DR PRINTS; PR00009; EGFTGF.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 DR NON TER 1 1
 FT NON TER 239 239
 SQ SEQUENCE 239 AA; 27024 MW; 49B7FDFDC91925B7 CRC64;

Query Match 17.2%; Score 224.5; DB 2; Length 239;
 Best Local Similarity 32.4%; Pred. No. 2.1e-09;
 Matches 68; Conservative 33; Mismatches 86; Indels 23; Gaps 8;

QY 45 SADGEFV---TSRSEMSGSEISPVSEMPSSSEPSGADYVSEY--DNEPQIPGIV 98
 DB 45 SLDRFESGKPRITVNLHTSEKNTITGQADNGSVGD-DYEDYEDYELNP----- 99
 QY 99 DSDVRVEQVYKPPQNTSESENTSDPKRRKKGKGNKRRNRKKNPCNAFONFCIHGE 158
 DB 100 ---RVAFSTKP--KHPSAMPTEKTKKRRKKGK-GKGRGN--KNLCKEYKDCIHGV 150
 QY 159 CKYIEHLAVTCKCOOEYFGERGCKSKMTHMIDSSLSKIALAALAFMSAVILTA 218
 DB 151 CHYQDLRTHSCVCGHGYSGERCHVFTLVGKEEDORYSRITLAVIAVLVS---LMCLAV 207
 QY 219 ITVQLRQYVRKYEGEAERKTLROENGAV 248
 DB 208 IAILALRYHKDDADVSESEKVKLEATSV 237

RESULT 11

Query Match	Best Local Similarity	Score	DB 1	Length	DB 2	DB 3	DB 4
Matches	55; Conservative	23; Mismatches	59; Indels	13; Gaps			
Qy 100	DSRVRYQVYVKKPQNTSESNISDKRRKKKGGKGNKRNRRKKRNPCNAEPQNFCHGRC	159					
Db 70	DLIRKTLTSSKPOLALAT--PWEHEGKRKKKKGGGLG-----KKRDPCLARKYDFCHGRC	121					
Qy 160	KYIEHLEAVTCKQCOQGYFGERGCKSMKTHMSIDLSKIALAALAFMSAVITLAVAVI	219					
Db 122	KYVEKELRAPSCTCHPGYHGERCHGLSLPVENRRLTYHDTTLIAVAVVLLSSVCL---LVI	178					
Qy 220	TVQLARQVYRK--VEGEAEERKKLROENGN	247					
Db 179	VGLLMFRYHRRGGYDVENEERKVLKGMTNSH	208					
RESULT 13							
HGGF_PIG							
ID_HBGP_PIG	STANDARD;	PRT;	208 AA.				
AC Q01580							
DT 01-FEB-1996	(Rel. 33, Created)						
DT 30-MAY-2000	(Rel. 39, Last sequence update)						
DT 05-JUL-2004	(Rel. 44, Last annotation update)						
DE	Heparin-binding EGF-like growth factor precursor (HB-BGF) (HBEGF).						
GN	Name=DR; Synonyms=HBEGFL, Sus scrofa (pig).						
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.						
NCBI	TaxID=9623;						


```

RN [1]
RP SEQUENCE FROM N.A.
RA Pascall J.C.
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RM [2]
RP SEQUENCE OF 99-182 FROM N.A.
RC STRAIN=large white; TISSUE=Heart;
RX MEDLINE=93d75016; PubMed=1445231;
RA Vaughan T.J., Pascall J.C., Brown K.D.;
RT "Tissue distribution of mRNA for heparin-binding epidermal growth
   factor."
RL Biochem. J. 287:681-684(1992).
CC -1- FUNCTION: May be involved in macrophage-mediated cellular
   proliferation. It is mitogenic for fibroblasts and smooth muscle
   but not endothelial cells. It is able to bind EGF receptors with
   higher affinity than EGF itself and is a far more potent mitogen
   for smooth muscle cells than EGF (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Mature HB-EGF is
   released into the extracellular space and probably binds to a
   receptor (By similarity).
CC -1- TISSUE SPECIFICITY: Macrophages, midbrain, cerebellum,
   hypothalamus, cerebral cortex, bulbourethral gland, lung, heart
   ventricle, kidney, skin, prostate, seminal vesicle, testis; at low
   levels in lymph node, thymus, spleen; not detected in pituitary,
   olfactory bulb, thyroid, duodenum, pancreas, liver, submaxillary
   gland.
CC -1- PTM: O-glycosylated (Probable).
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL, Y15731, CA475740.1, -
DR EMBL, X67295, CA47709.1, -
DR PIR, S27162, S27162.
DR HSSP, Q99075, 1XDT.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR006209; EGF_like.
DR Pfam, PF00008; EGF_1.
DR PROSITE, PS00022; EGF_1; 1.
DR PROSITE, PS01186; EGF_2; 1.
DR PROSITE, PS50026; EGF_3; 1.
KM EGF-like domain; Glycoprotein; Growth factor; Heparin-binding; Signal;
   Transmembrane.
FT SIGNAL 1 23 Potential.
FT PROPEP 24 62 By similarity.
FT CHAIN 63 148 Heparin-binding EGF-like growth factor.
FT PROPEP 149 208 C-terminal (Potential).
FT DOMAIN 24 161 Extracellular (Potential).
FT TRANSEM 162 182 Potential.
FT DOMAIN 183 208 Cytoplasmic (Potential).
FT DOMAIN 104 144 EGF-like.
FT DISULFID 108 121 By similarity.
FT DISULFID 116 132 By similarity.
FT DISULFID 134 143 By similarity.
FT CARBOHYD 85 85 O-linked (GalNAc...) (By similarity).
SQ SEQUENCE 208 AA; 22866 MW; 0A7DA97AB30C8967 CRC64;

Query Match 15.9%; Score 208.5; DB 1; Length 208;
Best Local Similarity 32.0%; Pred. No. 3e-08;
Matches 62; Conservative 30; Mismatches 75; Indels 27; Gaps 7;

```

```

Db 80 P0ALATPSK--EERGRKRRKGGKGLG-----KKRDPCLRYKDFCIHGECYIKELRAPS 131
Qy 170 CKCOQYFGRGCEKSMKTHSMIDSLSKALAAAFSAVILTAVALITVQLRQYR 229
   |||||
Db 132 CICHPGYHGERCHGLSLPVRNRLTYDHTTILAVAVVAVSLVCL--LVIVGLMFRYHR 188
   |||||

Qy 230 K--YEGAEERKTL 241
   |||||
Db 189 RGGDVENEKRVKL 202

RESULT 14
ID Q9W7C5 PRELIMINARY; PRT; 212 AA.
AC Q9W7C5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Heparin-binding EGF-like growth factor.
GN Name=HB-EGF;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=bred White Leghorn;
RX MEDLINE=99254107; PubMed=10318950; DOI=10.1073/pnas.96.10.5716;
RA Fu S.L., Bortolli I., Goller M., Vogt P.K.;
RT "Heparin-binding epidermal growth factor-like growth factor, a v-Jun
   target gene, induces oncogenic transformation."
   Proc. Natl. Acad. Sci. U.S.A. 96:5716-5721(1999).
RM [2]
RP SEQUENCE FROM N.A.
RC STRAIN=bred White Leghorn;
RL Fu S.L., Bortolli I., Goller M., Vogt P.K.;
RM Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
DR EMBL, AF131224, AAD29416.1, -
DR HSSP, Q99075, 1XDT.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR006210; IEGF.
DR Pfam, PF00008; EGF_1.
DR SMART, SM00181; EGF_1.
DR PROSITE, PS00022; EGF_1; UNKNOWN_1.
DR PROSITE, PS01186; EGF_2; 1.
DR PROSITE, PS50026; EGF_3; 1.
KM EGF-like domain.
SQ SEQUENCE 212 AA; 22541 MW; E82A8D08F5297183 CRC64;

Query Match 15.2%; Score 198.5; DB 2; Length 212;
Best Local Similarity 33.3%; Pred. No. 1.8e-07;
Matches 50; Conservative 24; Mismatches 61; Indels 15; Gaps 3;

```

```

Db 95 GYIVDSVRVEYVKKPPQNTSENTSDPKRRKKGGKGNRRNRKKNPCNAEPQFNC 154
   |||||
Qy 69 GDALSELPRVAFPSKGGPVT-----PKKGNKRRKRGKGLKRRDPCLRYKDFC 120
   |||||
Db 121 IHGECYIHELGPSCICQGYHGERCHGLLVEHPSPYDHTTILAVAVVAVSLSLCV 180
   |||||
Qy 155 IHGECYIHELVAITCKCOQYFGRGCEKSMKTH---SMIDSLSKALAAAFSAV 211
   |||||
Db 212 ILTAVALITVQLRQYRKYEGAEERKTL 241
   |||||
Db 181 IITALLMFRCHKR-----GYVDVENEKRVKL 206

RESULT 15
ID Q9BG62 PRELIMINARY; PRT; 132 AA.
AC Q9BG62;

```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
 DE Hepatin-binding epidermal growth factor (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
 RA Terens F., Fischer B.,
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 DR EMBL; AF333182; AAK14375.1; -
 DR HSSP; O99075; 1XDT.
 DR InterPro; IPR001336; EGF_1.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR006209; EGF_like.
 DR Pfam; PF00008; EGF_1.
 DR PRINTS; PR00009; EGF_TGF.
 DR SMART; SM00181; EGF_1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS50026; EGF_3; 1.
 KW EGF-like domain.
 FT NON_TER 1.
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 14628 MW; 6FSACCO7E6A0F4E CRC64;

Query Match 14.9%; Score 195; DB 2; Length 132;
 Best Local Similarity 34.9%; Pred. No. 1.8e-07;
 Matches 44; Conservative 22; Mismatches 52; Indels 8; Gaps 2;

QY 91 PQLPGYIVDSVRVEQYVKKPPQNTSESENTSDKKRKKGKNGKRNRRKKKNPCNAEF 150
 Db 15 PEVDDLEADLYRAAFSSKPOLATPSK--EERGRKKRKGGLG-----KKRDPCLARKY 66
 QY 151 QNFCHGCKYIEHLEAVTCKCOQEVFERCGEKSMKTHSMIDSLKIALAIAAFMSA 210
 Db 67 KDFCHGCKYIKELRAPSCICHGYSGERCHGSLPEVNRLYTYDHTTILAVAVVSS 126
 QY 211 VILTAV 216
 Db 127 VCLLVI 132

Search completed: February 26, 2005, 16:44:14
 Job time : 176 secs